Is there a Carbon Premium? Exploring Methodological Sensitivity and the Puzzle of Inconsistent Results

by

Floor van der Sanden
(SNR: 2041769)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Quantitative Finance and Actuarial Science

Tilburg School of Economics and Management
Tilburg University

Supervised by: Prof. Christoph Hambel
Second reader: Prof. Bas Werker

April 18, 2024

Contents

Abstract

Climate change presents substantial global challenges, prompting heightened awareness and action. Consequently, beyond the physical impacts of climate change, policy regulations aimed at climate change mitigation, along with shifts in consumer and investor attitudes, are introducing additional risks on the path toward a greener economy. An important area of inquiry is the effect of these risks on firm valuation, particularly the differential impact on polluting versus environmentally friendly firms. Theoretical perspectives offer varied conclusions regarding the correlation between a company's environmental performance and its financial performance, a diversity mirrored in empirical observations. Leveraging a dataset covering over 10,000 publicly listed firms across 90 countries from 2007 to 2023, we investigate the influence of methodological choices on research outcomes concerning the relationship between a firm's environmental and financial performance. Our primary finding provides evidence supporting a negative carbon premium, indicating that environmentally responsible firms tend to outperform their less environmentally conscious counterparts. However, we demonstrate that this conclusion is contingent upon several critical factors, including the treatment of unscaled emissions, the inclusion of vendor-estimated emissions, temporal considerations regarding emissions and accounting data, and, to some extent, the analytical framework employed, distinguishing between panel regression analysis and portfolio analysis. Our findings emphasize the necessity for cautious interpretation of previous research outcomes concerning the association between a company's environmental performance and its financial results.

Keywords: carbon premium, green outperformance, carbon emissions, stock returns, transition risk, climate change, climate finance

Acknowledgements

The main aim of this thesis has been to conduct a relevant analysis on the relationship between stock returns and firm's emissions, addressing a gap in the literature regarding the sensitivity of results to methodological choices. Despite encountering challenges and occasional frustrations, the process of writing this thesis has been deeply rewarding and a significant highlight of my academic pursuits.

I wish to express my gratitude to the individuals and institutions whose support and guidance were instrumental in the completion of this thesis. First, I am indebted to the academic staff of Tilburg University for imparting invaluable tools and knowledge throughout my academic journey. Their contributions have not only facilitated the writing of this thesis but will also remain significant in shaping my future career. In particular, I extend my appreciation to Prof. Christoph Hambel for his invaluable guidance, constructive feedback, and encouragement during the thesis writing process. Additionally, I am grateful to PwC – The Netherlands for providing access to the LSEG Workspace database, which enabled the expansion of my analysis across 90 countries

I extend my special appreciation to my friends, who have played diverse roles in my academic journey. Some have been invaluable partners in group projects, while others have provided welcome distractions, enriching student life. Particularly, Paul has been a consistent source of support, patiently listening to and engaging with my endless voice notes over the years. Lastly, heartfelt thanks are due to my family for their unwavering encouragement and unconditional support. Though my name appears on the diploma, theirs should rightfully be alongside mine.

Introduction

The ramifications of climate change spare no corner of the globe, manifesting in rising temperatures that fuel environmental degradation, natural disasters, economic upheaval, and societal unrest.[1] Recognizing the severity of these challenges, governments, institutions, and the global public are increasingly acknowledging the urgent need for action (IPCC, 2018). Consequently, beyond the physical effects of climate change, policy regulations aimed at mitigating its impact and shifts in consumer and investor attitudes are introducing additional risks on the path toward a greener economy. An important inquiry in this context is how these risks, emerging from uncertainties surrounding the realization and timing of the transition to a greener economy, influence firm valuation, and in particular, how this effect differs for polluting versus more environmentally friendly firms.

The intricate interplay between climate change and the economy is not new and has been a subject of study since the 1970s, with the groundbreaking work of Nobel laureate Nordhaus (2019). However, financial aspects of climate change economics have gained significant importance in recent years, particularly amidst escalating climate impacts and intensified mitigation efforts (Friede et al., 2015). The focus of our study on the valuation of environmentally friendly (green) versus environmentally detrimental (brown) firms has sparked substantial debate within climate finance research, the investment industry, and policy-making.

Understanding the connection between a firm's environmental performance and its financial performance holds significant importance for various stakeholders. For investors, this understanding enables informed decision-making aligned with sustainability goals. Yet, significance extends beyond investor interests, encompassing broader research objectives and potential societal benefits. Specifically, insights into the impact of firms' environmental performance on financial performance can guide the development of policies promoting sustainable practices in the business sector, thereby aiding in climate change mitigation efforts.

Zooming out to a broader perspective, studying climate change through the lens of financial economics is essential. Financial markets, crucial for risk-sharing, offer avenues for transferring the aggregate risk of climate change. Despite the pervasive nature of climate change risk, the diversity in exposure, adaptability, and risk tolerance across firms and regions provides opportunities for effective risk-sharing. Moreover, investors with varying risk profiles are positioned differently to bear and manage climate-related risks (Giglio et al., 2021).

Theoretical arguments propose divergent conclusions regarding the link between firm environmental performance and financial performance. The first argument centers on transition risk from policy regulations, potentially resulting in a 'carbon premium' for brown firms, as they face disproportionate impacts from emission-limiting policies (e.g., Bolton and Kacperczyk,

[1] For more information on the impact of rising temperatures and climate change in general, we refer to `https://www.un.org/en/climatechange`.

2021b, 2023).[2] Conversely, the taste argument contends that shifts in consumer and investor preferences could lead to green assets outperforming brown ones, despite lower expected returns (Pastor et al., 2019, 2022).

Empirical research mirrors this theoretical ambiguity, with findings varying between brown stocks outperforming green stocks and vice versa, or yielding mixed results. Notably, BK1 and BK2 discover that brown stocks outperform green stocks both in the US and globally, using emissions levels as a metric for a firm's environmental friendliness. Similar findings are reported by Hsu et al. (2023) for their US sample, utilizing a combination of emissions types as a measure of eco-friendliness. Trinks et al. (2022) additionally identify green stocks outperforming brown stocks, attributing this premium to uncertain firm-level impacts from future regulatory and market initiatives addressing climate change.

The landscape of climate finance research has seen numerous studies challenging the findings of influential papers, particularly those by BK. For instance, Bauer et al. (2022) present evidence contrary to the existence of a carbon premium, addressing contradictory findings on the link between emissions and stock returns. Additionally, Pastor et al. (2019) suggest that the outperformance of green stocks depends on heightened climate concerns. Moreover, debates around BK's studies are raised by Zhang (2024) and Aswani et al. (2023), with the former contending that the carbon premium results from forward-looking bias and the latter associating it with vendor-estimated emissions rather than actual emissions, advocating for emissions intensity as a more appropriate metric, contrary to BK.

The inconclusive empirical landscape regarding the link between a firm's environmental performance and its financial performance underscores the substantial impact of methodological choices on research findings. In this study, we undertake a comprehensive analysis of the link between emissions data and stock returns in an international sample of more than 10,000 listed companies spanning 90 countries, over a period ranging from 2007 to the beginning of 2023. Encompassing most of the companies for which our data vendor, LSEG Workspace, provides emissions data, our analysis aims to address the sensitivity of findings on the link between firms' environmental performance and financial performance, to methodological choices. As a measure of a firm's environmental performance, we rely on the sum of scope 1 and scope 2 carbon emissions.[3] We assess both emissions levels and emissions intensity, the latter being emissions normalized by a firm's revenues.

We differentiate between US and global samples, conduct panel regressions and portfolio analyses, assess reported versus vendor-estimated emissions, and investigate the impact of lagging emissions or accounting data. When conducting panel regressions, we regress returns on either log emissions levels or emissions intensity, controlling for firm-specific characteristics and fixed effects. In portfolio analyses, we create a Green-Minus-Brown (GMB) spread by taking a long position in the portfolio comprising the quintile of greenest firms, and a short position in the quintile of brownest firms. Subsequently, we regress the GMB spread on a constant and factors from the Fama-French 5-factor model and the momentum factor (Carhart, 2012; Fama

[2]Given the high number of times we cite Bolton and Kacperczyk (2021b) and Bolton and Kacperczyk (2023), henceforth, we will denote the first paper as BK1 and the second as BK2. Should we need to reference both papers concurrently, we will cite them as BK.

[3]Carbon emissions are typically categorized into three distinct scopes according to the GHG Protocol (WBCSD and WRI, 2004). Scope 1 includes direct emissions from sources owned or controlled by a firm, scope 2 comprises indirect emissions from purchased electricity consumption, and scope 3 encompasses all other indirect emissions within a firm's value chain.

and French, 1993, 2015).

Our primary finding challenges the notion of a positive carbon premium. In the US, green outperformance is evident in panel regressions using log emissions. For the sample encompassing reported and vendor-estimated emissions, green outperformance averages between 1.20% and 1.30% points monthly, depending on the specification. However, this effect stems from vendor-estimated emissions levels and diminishes when scaling emissions by revenues. While green outperformance persists for vendor-estimated emissions when lagging emissions levels by 6 or 10 months and accounting data by 6 months, it is not robust enough in the sample containing reported and vendor-estimated emissions. Portfolio analyses also indicate green outperformance in the US, not solely driven by vendor-estimated emissions and resilient to lagging emissions data by 6 or 10 months. Alphas range from approximately 1.02% to 1.30% points when accounting for common risk factors. Ultimately, the green outperformance remains unexplained by abrupt shifts in climate change concerns, as evidenced by the inclusion of a proxy for these shocks constructed by Ardia et al. (2022) in our regressions.

Similar patterns are observed in the global sample, with green outperformance in panel regressions primarily driven by vendor-estimated emissions levels and diminishing when scaling emissions by revenues. While green outperformance persists for vendor-estimated emissions with lagging emissions levels by 6 or 10 months and accounting data by 6 months, it is not robust enough in the sample containing reported and vendor-estimated emissions. However, in portfolio analyses, green outperformance is also driven by vendor-estimated emissions levels, contrary to the US findings, and diminishes when scaling emissions levels by revenues. Moreover, while heightened climate change concerns do not fully explain green outperformance, their inclusion weakens the link between vendor-estimated emissions and stock returns in portfolio analyses, suggesting a more intricate global relationship. Notably, green outperformance in the global sample is less pronounced than in the US sample.

This study contributes to the existing literature in three main ways. Firstly, our main contribution is to emphasize the sensitivity of results regarding the link between a firm's emissions and returns to methodological choices. Our foremost finding suggests evidence in support of a negative carbon premium, indicating that green firms tend to outperform brown firms. However, we demonstrate that this result depends upon several critical factors, including the treatment of unscaled emissions, the inclusion of vendor-estimated emissions, temporal considerations regarding emissions and accounting data, and, to some extent, the analytical framework employed, distinguishing between panel regression analysis and portfolio analysis.

Specifically, we find that the negative carbon premium disappears when (log) emissions levels are scaled by revenues, both in the US data and the global sample, across all models. We highlight that variations in a firm's size, as previously noted by Zhang (2024), explain a significant portion of the variation in emissions levels. Therefore, our results indicate that the association between emissions levels and stock returns may reflect a relationship between firm fundamentals and stock returns or may be influenced by multicollinearity issues arising from including emissions levels and the logarithm of market capitalization as regressors within a single model. Furthermore, except for portfolio analyses in the US sample, we find that the negative carbon premium observed in the sample including reported and vendor-estimated emissions is primarily driven by vendor-estimated emissions. Finally, lagging emissions levels attenuate the link between emissions and stock returns in the sample including reported and vendor-estimated emissions, particularly in panel regressions, suggesting the importance of aligning data with investor knowledge.

Our findings highlight the necessity for cautious interpretation of previous research outcomes. Our study aligns closely with recent works by Zhang (2024) and Aswani et al. (2023), who similarly scrutinize factors in the existing literature on the relationship between a firm's environmental and financial performance. While previous studies have addressed such factors, our approach distinguishes itself by considering multiple variables comprehensively, contrasting with the more focused analyses of the aforementioned authors.

BK, alongside Zhang (2024) and Aswani et al. (2023), initially find indications of a positive carbon premium when examining emissions levels from S&P Trucost. This differs from our findings of green outperformance, aligning with Bauer et al. (2022), who also use data from LSEG Workspace. This discrepancy prompts inquiry into whether research outcomes depend on firm coverage and the choice of data vendor. Emissions data emerges as a crucial factor potentially contributing to divergent results across vendors, leading to our second key contribution.

As a second contribution, our paper stands out for its thorough analysis of the quality and reliability of carbon emissions data supplied by LSEG Workspace. While Bajic et al. (2023) have systematically assessed emissions data quality, prior studies using emissions data have not comprehensively investigated its reliability. Notably, influential studies by BK, Aswani et al. (2023), and Bauer et al. (2022) use emissions data without addressing its reliability. Similarly, in other areas of climate finance, researchers often overlook these issues. For example, Matsumura et al. (2014) explore the effects of carbon emissions and voluntary disclosure on firm value, while Kabir et al. (2021) examine the impact of carbon emissions on firm default risk. Though we focus on emissions data from LSEG Workspace, discrepancies uncovered may indicate similar concerns with emissions data from other providers, potentially impacting broader insights into the financial implications of emissions.

Third, our approach stands out in the literature due to its relatively large sample, covering a period marked by various highly uncertain events, including the 2008 economic crisis, the COVID-19 pandemic, the Ukraine war and energy crisis, and the Israeli–Palestinian conflict. To assess the impact of climate change concerns stemming from these events, we incorporate a factor indicating climate change news shocks, as constructed by Ardia et al. (2022). However, in our monthly analyses, we do not find evidence suggesting that the observed green outperformance can be attributed to sudden increases in climate change concerns.

The subsequent sections of this paper are structured as follows. Chapter 1 provides an overview of the relevant literature, covering climate change risks, theoretical hypotheses regarding the relationship between a firm's environmental and financial performance, and the empirical literature, emphasizing methodological differences. Chapter 2 outlines the dataset and our data cleaning process, focusing on refining emissions data. In Chapter 3, we offer an examination of the dataset, exploring variations in emissions levels and intensity across different countries, comparing reported emissions with vendor-estimated data, and discussing the financial information captured by emissions data. Chapter 4 elaborates on our methodological framework, encompassing both panel regressions and portfolio analyses. Our findings are presented in Chapter 5, followed by the conclusion in Chapter 6, and suggestions for future research directions in Chapter 7.

Chapter 1

Literature review

The inconclusive literature on the relationship between a firm's environmental performance and its stock returns calls for a detailed review. This chapter first introduces the two main prevalent mechanisms that link a firm's environmental performance to its financial performance, starting with an introduction of the different risks induced by climate change.

Second, this chapter is dedicated to providing a comprehensive overview and comparative analysis of various empirical studies investigating on this link, to identify potential factors contributing to the divergent results prevalent in the literature. We will examine these studies, emphasizing their methodologies, findings, and acknowledged limitations. Subsequently, we will identify the differences in methodologies employed when comparing the papers under discussion.

1.1 Induced risk from climate change

The literature on the financial applications of carbon emissions suggests two main mechanisms through which a firm's environmental impact can influence its financial performance. Before discussing these mechanisms, we first introduce the most prevalent risks associated to climate change, forming the basis of these mechanisms.

Climate change may cause severe risks for firms, consumers and investors. There is broad consensus that these risks can be classified into two categories, physical risk and transition risk. First, physical risk relates to the physical impact of a changing climate, including the potential destruction of assets or the disruption of operations, trade routes, supply chains, and markets. Examples encompass gradual processes, such as the rising sea levels, as well as sudden occurrences including floods and heatwaves. These events are affected by climate change, contributing to alterations in their frequency and intensity (e.g., Sakhel, 2017; NGFS, 2022).

The second risk, carbon transition risk, is particularly pertinent to this paper. In this paper, we define transition risk in its broadest sense, similar as to BK, with the risk being an amalgamation of a wide range of factors. Existing literature outlines three primary drivers for transition risk (e.g., Campiglio and van der Ploeg, 2022). The first driver, which is the main focus of most of the research on this topic, is linked to the implementation of climate mitigation policies, with governments and institutions progressively adopting regulations to achieve carbon neutrality by 2050. This ambitious initiative aims to prevent an increase in average temperatures beyond 1.5 °C compared to pre-industrial levels, a threshold deemed critical to human survival (NGFS, 2022). Given the uncertainties surrounding the pace of adaptation to carbon neutrality, firms face risks in adjusting to these measures.

The second form of transition risk is centered around sudden changing consumers' and investors' beliefs and expectations concerning environmental issues. More specifically, investors may not solely concentrate on financial metrics such as returns, but may also take into account non-financial factors, such as a firm's role in climate change, when making investment decisions (e.g., Pastor et al., 2019, 2022). Potential drivers of a rapid change include dramatic climate-related events or exposure to news related to climate change.[1] Finally, we note that regulatory risk may also alter investor's beliefs, since investors face evolving beliefs and expectations related to the transition to a net-zero world.

The third form of transition risk concerns unanticipated or rapid improvements in technology. These can make carbon-intensive firms obsolete, impacting share prices. Alternatively, emerging negative emissions technologies may sustain such firms (e.g., Campiglio and van der Ploeg, 2022; Hambel and van der Ploeg, 2024).[2]

Both physical and transition risk should be accounted for by investors, but measuring physical climate risk is challenging due to its complexity, as noted by Loyson et al. (2023). Therefore, our focus is solely on transition risk, stemming from policy changes or shifts in investor and consumer beliefs. Crucial is that policy transition risk materializes independently of any potential future physical damages resulting from climate change (Meinerding and Zhang, 2023). Furthermore, the risks may increase as the deadline to carbon neutrality approaches, especially if companies do not adapt to regulations (Bolton and Kacperczyk, 2021b). Additionally, climate transition risk can transcend borders, impacting companies globally as investors react to international events and policy changes (Yang et al., 2024).

1.2 Hypotheses underlying a carbon premium

In the literature examining the financial impact of carbon emissions, two main mechanisms are proposed linking a firm's environmental performance to its financial outcomes. The first mechanism emphasizes the uncertain impacts at the firm level resulting from future regulatory and market actions addressing climate risks. Firms with high carbon emissions are expected to be disproportionately affected, leading to higher regulatory transition risks, and investors may demand higher expected returns to compensate for this risk (e.g., Bauer et al., 2022; Bolton and Kacperczyk, 2023; Pastor et al., 2019). This phenomenon is commonly known as the 'carbon risk premium' or 'carbon premium'.[3]

However, in recent years, the investment industry has promoted the idea of potentially higher average returns for green firms compared to brown firms, under the rubric 'doing well by doing

[1] Noteworthy examples include studies indicating that personal experiences with extreme weather events shape public views on climate issues, leading to increased support for government climate action (Harvard, NPR and RWJF, 2022). Exposure to authoritative reports on climate change, such as the IPCC special report, has also been associated with heightened climate change concern (Ogunbode et al., 2019).

[2] Negative emissions technologies (NETs) are technologies that remove CO_2 emissions from the atmosphere. Examples of current technologies include direct air capture of CO_2 via artificial trees and Augmented ocean disposal, using lime in oceans to trap CO_2 in a stable, dissolved inorganic form (McGlashan et al., 2012). However, existing NETS are not (yet) competitive given high costs and political economy challenges (Honegger and Reiner, 2017). Nevertheless, breakthroughs in technologies may make NETs worthwhile in the future. Due to the uncertainty involved in research developments, NETs add to transition risk.

[3] We note that a premium can be both positive and negative. However, the 'carbon premium' is used to describe a positive premium. To avoid ambiguity, we will often refer to a 'positive carbon premium' or a 'negative carbon premium'.

good' (Karnani, 2011). Given that green stocks are less exposed to regulatory transition risk and are favored by investors who prioritize non-financial characteristics, higher returns for green stocks contradict the predictions based on asset pricing theory. However, asset pricing theory primarily focuses on expected returns, whereas we observe realized returns. According to Pastor et al. (2019, 2022), these returns can significantly differ in the face of unexpected shifts in preferences or risk perceptions.

Notably, Pastor et al. (2019) develop a theoretical framework to explain the gap between expected and realized returns. In this framework, referred to as the PST framework in this paper, consumer preferences can suddenly shift away from less eco-friendly firms towards greener alternatives, for example when climate concerns unexpectedly rise. This shift negatively impacts the profitability of polluting firms, reducing their market value, while the opposite occurs for environmentally responsible companies. Furthermore, heightened climate concerns result in investors finding greater value in investments in environmentally sustainable firms, driven by either intrinsic motivations or societal pressure to disinvest from environmentally harmful companies. This increased demand prompts selling pressure and higher discount rates, causing less eco-friendly firms to depreciate in value and green firms to appreciate. Should these unforeseen shifts be substantial, assets aligned with green initiatives may outperform those associated with brown sectors, even if they exhibit lower anticipated returns. This theory challenges the notion that transition risk favors brown firms over green ones, suggesting that a sudden, widespread increase in transition risk would boost the valuation of green firms over brown ones.

It is crucial to highlight that the theoretical framework proposed by Pastor et al. (2019) predicts that green assets will underperform brown assets over a sufficiently long observation period, aligning with asset pricing theory. However, the model serves as a cautionary reminder of why using realized returns as a substitute for expected returns can be deceptive when sudden unexpected shifts in preferences or risk perceptions occur among consumers and investors.

1.3 Prior empirical evidence

In recent years, empirical research on the connection between a firm's environmental performance and its financial performance has grown substantially, resulting in diverse findings. Studies either affirm, refute, or present mixed results regarding the existence of a carbon premium and the sign of it. We highlight influential papers in this area, acknowledging that our analysis does not encompass all literature on this topic. Initially, we review papers supporting a positive carbon premium, followed by those presenting contrasting findings, including evidence of a negative premium or no premium at all.

Two of the most impactful articles in this particular field are the works authored by Bolton and Kacperczyk in 2021, and subsequently, in 2023 (Bolton and Kacperczyk, 2021b, 2023), denoted by BK1 and BK2 in this paper, as introduced earlier.[4] In their first paper, the authors focus mainly on US data while they extend their analysis in the second paper, investigating the impact of firm-level carbon emissions on stock returns in conjunction with country-specific characteristics reflecting a nation's progress in energy transition. Regression analyses confirm a

[4] While the two papers share similar findings and methodologies, with the work by Bolton and Kacperczyk (2023) essentially building upon the foundation laid by the paper from Bolton and Kacperczyk (2021b), we maintain separate references to these papers to eliminate any potential confusion.

statistically significant carbon premium, consistent across emissions scopes and regions, both in the short term and the long term.

Another study supporting the presence of a carbon premium is authored by Hsu et al. (2023). In their research focusing on the US, they explore the link between industrial pollution and stock returns. They find statistically significant evidence of brown firms outperforming green firms. The authors explain this 'pollution premium' through higher regulatory risks faced by polluting firms and provide empirical support for this hypothesis.

We conclude the discussion on the carbon premium with Trinks et al. (2022), who differentiate between systematic and non-systematic risk. Initially, the authors suggest that investors screening out high-carbon assets may lead to green outperformance. However, their portfolio analyses do not support this claim. To examine systematic risk, which they argue to be the result of regulatory risk and direct climate change effects, Trinks et al. (2022) utilize panel regression techniques. They find a consistently positive effect of carbon intensity on the Cost of Equity Capital (CoE), particularly pronounced in high-emission sectors, European Union nations, and firms subject to carbon pricing regulations.

After BK's influential papers supporting the carbon premium, numerous studies have emerged to critique these findings. We provide an overview of significant contributions to this discourse. First, Bauer et al. (2022) investigate carbon emissions' influence on stock returns, expanding their analysis from the US to include G7 countries[5]. Their preferred method involves constructing portfolios of green and brown firms based on emissions levels or intensities, resulting in evidence in favor of green outperformance. Despite additionally employing panel regressions and controlling for fixed effects and firm characteristics, Bauer et al. (2022) find no evidence of a positive carbon premium.

Another study by Pastor et al. (2022) examines realized returns versus expected returns' potential issues, revealing green stocks' outperformance tied to increased climate concerns. This result cautions against expecting consistently high returns in the future. It is consistent with the model proposed by Pastor et al. (2019), the PST framework as discussed in Section 1.2, emphasizing the role of unexpected shifts in climate concerns in driving stock performance.

Furthermore, Zhang (2024) critically examines methodologies used by BK and related studies, reevaluating BK's findings on the carbon premium in the United States and globally. Emphasizing the importance of lagging carbon emissions to avoid a forward-looking bias[6], the author employs portfolio analysis and discovers evidence contradicting the existence of a carbon premium. Additionally, Zhang (2024) benchmarks against BK, demonstrating that, after controlling for sales during the emission period, the positive association between total emissions, emission growth, and stock returns diminishes. The author concludes that the alleged carbon premium is solely tied to strong firm performance during the emission period and does not represent a risk premium associated with transition risk. Furthermore, Zhang (2024) notes that carbon intensity effectively purges sales information, removing the forward-looking bias induced by sales information included in emissions data.

Aswani et al. (2023) scrutinize BK's methods, finding two key points. Firstly, they distinguish between firms reporting actual emissions and those with estimated values. This segregation reveals that the relationship between stock returns and emissions in the U.S., as demonstrated

[5]The G7 countries comprise the United States, Canada, France, Germany, Italy, Japan, and the United Kingdom.
[6]The forward-looking bias is discussed in greater detail in Subsection A.1.5 of the Appendix.

by BK1, is predominantly associated with vendor-estimated emissions rather than firm-disclosed actual emissions. Secondly, they advocate for using emissions intensity over levels but find no evidence supporting a carbon premium in their analysis when focusing on this emissions intensity. They note these issues extend to EU data.

1.4 Identifying differences across methodologies

The above discussion highlights the conflicting outcomes observed in studies investigating the relationship between a firm's sustainability level and its stock returns. Our study serves as a commentary on the methodologies employed in this realm, aiming to uncover factors contributing to these discrepancies. We identify six key areas of divergence, including modeling approaches, greenness metrics, data sources and ESG scores, issues surrounding carbon emissions, timing disparities, and variations in time frames and regions considered. In this section, we only summarize the latter areas and we refer to Appendix section A.1 for a detailed discussion.

First, studies adopt a variety of modeling approaches, ranging from regression analyses to traditional empirical asset pricing methods. These methodologies differ fundamentally. While regression models can incorporate industry-fixed effects and firm-specific characteristics, traditional methods primarily control for common risk factors such as the Fama-French 5 factors (Fama and French, 1993, 2015).

Additionally, the choice of metrics for assessing a firm's environmental impact presents another area of divergence. Some studies focus on emissions levels, while others derive greenness scores from ESG datasets. Greenness scores have faced criticism due to inconsistencies, and alterations in ESG scores over time further complicate the landscape (e.g., Charlin et al., 2022; Berg et al., 2022; La Torre et al., 2020). Discrepancies also arise in the evaluation of carbon emissions data between different vendors, highlighting challenges in data reliability and interpretation Papadopoulos (2022).

Furthermore, challenges associated with emissions data, including inaccuracies and the inclusion of vendor-estimated emissions, complicate research findings (e.g., Aswani et al., 2023; Bajic et al., 2023). Specifically, data providers can provide estimated emissions based on their own models for companies that did not disclose emissions. The ongoing debate over emissions levels versus emissions intensity adds another layer of complexity, with studies showing differing results based on these methodological choices. Finally, disparities in time frames and regions studied further contribute to the sensitivity of research outcomes.

Empirically, the sensitivity of results to methodological choices underscores the need for careful consideration of these factors. While a compact overview of methodologies is provided in Table 1.1, a more detailed overview can be found in Table C.1 in the Appendix. In the subsequent chapters, we will explore these methodological considerations, providing further insights into the nuances of this relationship.

Table 1.1: Tabulated compact overview of the baseline methodologies employed in different studies investigating the influence of a company's environmental performance on its stock performance. For a more elaborate overview, see Table C.1 in the Appendix.

Paper	**Model**	**Greenness measure**	**Data vendor**	**Time frame**	**Geographic Scope**	**Results**
BK1 and BK2.	Regress environmental measures against monthly stock returns and control variables using pooled OLS.	Emissions levels, emissions growth.	Trucost, Factset.	2005 - 2017 / 2005-2018.	U.S. / 77 countries.	Brown stocks outperform green stocks.
Hsu et al. (2023).	Assess the value-weighted quintile return spread, using their emissions measure for sorting.	Aggregated emissions for all types of chemicals.	TRI, Comput-stat, CRSP.[7]	1991 - 2016.	U.S.	Brown stocks outperform green stocks, across all defined forms of emissions intensity.
Trinks et al. (2022).	Two separate models. Model 1: Assess the value-weighted decile return spread. Model 2: Regress environmental measures against monthly CoE and control variables using pooled OLS.	Emissions intensity and sector-adjusted carbon intensity.	Refinitiv.	2008 - 2016.	50 countries.	Brown stocks outperform green stocks and this effect primarily relates to carbon transition risk (systematic risk) and not to investor's preferences (non-systematic risk).
Bauer et al. (2022).	Baseline: Assess the value-weighted monthly quintile return spread, using their emissions measures for sorting. Secondary: regression analysis.	Emissions levels and emissions intensity.	Refinitiv.	January 2010 - December 2021.	G7 countries.	Green stocks outperform brown stocks, particularly considering emissions levels and emissions intensity as measures of greenness.
Pastor et al. (2019).	Assess the monthly value-weighted tertile return spread, using their greenness measure for sorting.	Unique environmental score.	MSCI.[8]	November 2012 - December 2020.	U.S.	Green stocks outperform brown stocks but this outperformance likely reflects an unanticipated increase in environmental concerns.
Zhang (2024).	Baseline: assess the monthly value-weighted quintile return spread, using their emissions measure for sorting. Secondary: pooled OLS, similar as to BK.	Carbon intensity.	Same as BK.	June 2009 - December 2021.	U.S.	Green stocks outperform brown stocks. The look-ahead bias in the analysis of BK overstates the carbon premium.
Aswani et al. (2023).	Baseline: replicated version of BK1.	Emissions levels and emissions intensity.	Similar to BK1.	2005-2019.	U.S.	Brown stocks outperform green stocks when relying on emissions levels, but only when considering vendor-estimated emissions.

[7] TRI is short for Toxic Release Inventory. CRSP is short for Center for Research in Security Prices.
[8] MSCI is short for Morgan Stanley Capital International Index.

Chapter 2

Data and data cleaning

2.1 Firm selection

In conducting our analysis, we acquire firm-level financial data and emissions data from LSEG Workspace, formerly recognized as Refinitiv. LSEG Workspace's ESG database, previously known as Asset4, distinguishes itself as one of the most comprehensive in the industry. This recognition is attributed to its extensive database history, wide-coverage, diverse array of available variables, organizational capabilities, and specialized expertise.[1]

To select the data, we pre-select public firms with available ESG measures. LSEG Workspace allows the option to download firms on a regional or country-level basis.[2,3] We extract firm information from regions including the United States, Europe, Asia Pacific, North America, Latin America, Africa, and Middle East and North Africa (Mena). Our primary sample spans from January 2007 to January 2023. We restrict our sample from 2007 onwards since emission reporting is relatively poor before this time period (Meinerding and Zhang, 2023). Throughout our analysis, we rely on International Securities Identification Numbers (ISINs) to uniquely identify firms.

Acknowledging the ambiguity in defining countries within the Asia Pacific, Latin America, and Mena regions, our emphasis is on inclusivity. Rather than concentrating on distinct regions, our objective is to encompass a diverse array of firms from different countries. For an overview of the number of firms per region, please refer to the Appendix, Table C.2. For each selected firm, we retrieve financial, accounting, and emissions data.

[1] In recent times, LSEG Workspace has harmonized its ESG methodologies and data measurements with the Task Force on Climate-Related Financial Disclosures (TCFD) framework, a standardized framework for reporting. For a more in-depth understanding of the TCFD method, please refer to the TCFD report: `https://assets.bbhub.io/company/sites/60/2021/10/FINAL-2017-TCFD-Report.pdf`.

[2] The lists of firms per region include public firms active on the list's creation date, typically rearranged annually by LSEG Workspace.

[3] Since LSEG Workspace's emissions data can be updated over time, it is important to note for replication purposes, that we downloaded our final data over the course of December 2023.

2.2 Accounting and financial data

For each included firm, we collect monthly returns data, alongside accounting information. Although our primary analysis encompasses the period from 2007 onward, we acquire return data from 2006 to compute the momentum and volatility variables, which serve as control variables in the regression analyses. We come back to the control variables in Chapter 4. Nominal values are standardized in US dollars. However, we identify several firms for which certain variables' currencies could not be set to the US dollar. Consequently, we rectify such discrepancies by excluding these firms. The precise count of excluded companies per region due to this currency-related issue is explicated in Table C.2 in the Appendix.

Finally, we consolidate the data from different regions into a unified dataset and identify companies present in multiple regions based on unique ISIN numbers. Duplicate entries are limited to the US and North America datasets, leading to the removal of 3530 companies from the North American dataset that are also present in the US dataset. This results in 11537 unique companies. In addition to financial and accounting data, we also extract and subsequently amend emissions data. Next, we outline this process.

2.3 Emissions data

To measure a firm's environmental impact, we refrain from relying on ESG scores and instead direct our attention towards quantifiable emissions measured in tonnes of CO_2 equivalent (tCO_2e)[4], for the reasons as described in Section A.1 in the Appendix.

Our data retrieval process involves obtaining separate data on scope 1 and scope 2 emissions, as well as the aggregated total of scope 1 and scope 2 emissions, provided as a separate variable by LSEG Workspace.[5] It is worth noting that we do not focus on scope 3 emissions, for the reasons provided in Appendix section A.1.

In some of our analyses, we differentiate between vendor-estimated and reported emissions data to explore potential variations in the impact of vendor-estimated versus reported emissions on stock returns. As highlighted in Section A.1 in the Appendix, data providers may offer estimated emissions when a company fails to disclose such data. Despite the growing emphasis on ESG considerations, only half of all companies in LSEG Workspace's ESG coverage report on CO_2e emissions data. To bridge this void, the data vendor employs three distinct estimation techniques in case a firm does not report emissions.

LSEG Workspace provides two distinct types of aggregated emissions data. First, they report the variable 'Total CO2 Equivalent Emissions', and second, they report the variable 'Estimated Total CO2 Equivalent Emissions'. The former encompasses emissions solely when a firm reports them, whereas the latter incorporates both reported emissions and those estimated by LSEG Workspace. We retrieve both variables for further analysis in subsequent sections.

LSEG Workspace employs three estimation techniques in case a firm does not report emissions. These include 'CO2', 'Energy', and 'Median' models, which are sequentially applied until

[4]In particular, CO_2 equivalent emissions encompass carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), hydrofluorocarbons (HFCS), perfluorinated compound (PFCS), sulfur hexafluoride (SF_6) and nitrogen trifluoride (NF_3). LSEG Workspace follows greenhouse gas (GHG) protocol for all their emission classifications by type, see `https://ghgprotocol.org/`.

[5]We obtain both scopes separate as well the sum of the scopes for data cleaning purposes, as elucidated in Subsection 2.4.7

an estimate is obtained. Instances where none of these methods yield an estimate result in the absence of emissions data for the corresponding period and firm. The 'CO2' model uses the latest available CO_2e emissions values, adjusted for changes in revenues and employees. The 'Energy' model compares a firm's energy consumption with others in the industry to estimate emissions, considering revenues and employees. The 'Median' method assigns the industry's median emissions value, also considering revenues and employees. For a comprehensive understanding of the distinct estimation methodologies employed, we direct attention to the documentation provided by LSEG Workspace.[6] It is crucial to note that LSEG Workspace exclusively applies estimation techniques to the total of scope 1 and scope 2 emissions. Consequently, our analysis centers on the aggregate value of scope 1 and scope 2 emissions.

Emissions data are typically reported or estimated on an annual basis. However, our analysis is on a monthly basis. When retrieving monthly data, which is retrieved on the first of every month, LSEG Workspace sets emissions constant across a firm's fiscal year.

2.4 Data cleaning

As stated before, the primary objective of this paper extends beyond investigating the influence of a firm's emissions on returns. Our work also serves as a commentary on the methodological frameworks and data cleaning procedures employed in previous research concerning the association between a firm's emissions and returns. A pivotal facet of this evaluation lies in the data cleaning process. This section elucidates aspects that authors in previous research may have either overlooked or only partially considered regarding flaws or potential issues in raw data. Considering LSEG Workspace's recognized track record for the relatively precise reporting of financial and accounting data, this section places primary emphasis on the data cleaning process for emissions. The content of this section draws significantly from observations made by Bajic et al. (2023), who have provided insights into LSEG Workspace's emissions data. We direct attention to Figure B.1 for a visual representation of the data cleaning process, with a specific focus on cleaning emissions data as described in the subsequent subsections.

2.4.1 Harmonizing monthly analysis

In our analyses, we differentiate between estimated and reported emissions, relying on the estimation method variable provided. Typically, both emissions levels and the estimation method are reported annually. However, LSEG Workspace integrates emissions levels into their primary database, maintaining constant values across a firm's fiscal year, enabling the extraction of monthly data for emissions levels. In contrast, following an extensive dialogue with LSEG Workspace, we recognize that the estimation method has to be retrieved from another database and is only available on an annual basis.

To harmonize with the monthly frequency of our analysis, we incorporate the estimation method by presuming its constancy throughout a firm's fiscal year. This spans the 11 months preceding LSEG's reporting of the estimation method and the month of reporting. For a detailed view of the process of converting the yearly estimation method to monthly data, please refer to the Appendix, Section A.2.1.

[6] This information is no longer publicly available. Although we can access the documentation, we abstract from elaborating on the methodology for this reason.

2.4.2 Temporal alignment: lagging the estimation method

As discussed in the preceding section, the estimation method for a firm's emissions is obtained on an annual basis, necessitating an adjustment to synchronize with the monthly frequency of our analysis. It is imperative to note that a firm's fiscal year-end typically falls on the last day of the month, either the 30th or 31st. Consequently, the estimation method is reported on the 30th or 31st of a month. Conversely, financial, accounting, and emissions levels data are downloaded on the first day of each month.

To illustrate this temporal misalignment, consider firm X in January 2022. Our focus for financial and emissions data is on the date 01-01-2022, while the estimation method corresponds to 31-01-2022. Despite the matching month and year in both datasets, there exists an almost one-month time difference. In response, we address this incongruity by implementing a one-month lag on the estimation method. In the example, this involves aligning financial and emissions data on 01-01-2022 with the estimation method on 31-12-2021. For a comprehensive rationale behind this methodological choice, please consult the Appendix, Section A.2.2. Henceforth, references to the estimation method denote the one-month lagged estimation method. Since the estimation method is lagged by one month, there is no estimation method available for January 2007, the initial month in our sample. Thus, the sample period is limited to February 2007 through January 2023.

2.4.3 Comparing the implicit and explicit estimation method

As elucidated in Section A.2.2 in the Appendix, the variables 'Total CO2 Equivalent Emissions' and 'Estimated Total CO2 Equivalent Emissions' implicitly convey information about the estimation method, while the estimation method is also explicitly stated as a separate variable. We assess the extent of alignment between these pieces of information.

Given that the variable 'Total CO2 Equivalent Emissions' exclusively encompasses reported emissions, instances where the variable 'Total CO2 Equivalent Emissions' is filled for a company at a specific date should correspond to an estimation method of 'Reported'. Among the $519,499$ instances where the 'Total CO2 Equivalent Emissions' variable is filled, we identify $65,158$ cases where this correspondence is not met. Subsequently, we exclude these observations from the dataset.

Furthermore, when the variable 'Estimated Total CO2 Equivalent Emissions' is filled while 'Total CO2 Equivalent Emissions' is not, it implies that LSEG Workspace estimated the emission value. In such cases, the estimation method should be 'CO2', 'Median', or 'Energy'. The dataset contains $471,515$ instances where, based on the two emissions variables, we anticipate the estimation method variable to be one of these three methods. However, we identify $52,896$ cases where this expectation is not fulfilled, prompting the exclusion of these observations from the dataset. Finally, we eliminate observations where both variables on emissions are empty. There are no cases where the variable 'Estimated Total CO2 Equivalent Emissions' is empty while the variable 'Total CO2 Equivalent Emissions' is filled.

2.4.4 Comparing the two variables on emissions

In this paragraph, we conduct a detailed comparison between the variables 'Total CO2 Equivalent Emissions' and 'Estimated Total CO2 Equivalent Emissions', as previously introduced. The dataset comprises $454,341$ instances where the variable 'Total CO2 Equivalent Emissions' is filled and $872,960$ instances where 'Estimated Total CO2 Equivalent Emissions' is filled. In principle, all the required information on reported and vendor-estimated emissions should be captured by the variable 'Estimated Total CO2 Equivalent Emissions'. It is expected to reflect the reported value when emissions are reported and to equal the estimated value when LSEG Workspace provides an estimate. However, a meticulous comparison between the two variables enables the identification of potential discrepancies in emissions values.

Specifically, we undertake a comparison of the two variables when both are populated. Such instances indicate that the emissions value was reported by the firm, and both variables should exhibit identical values. We define the difference for a firm at a given date as the percentage difference, by dividing the absolute difference by the minimum of the two values. Specifically for firm i at time t, we define

$$\text{Difference}_{i,t} = \frac{\left|\text{Total CO2 Eq. Emissions}_{i,t} - \text{Estimated Total CO2 Eq. Emissions}_{i,t}\right|}{\min(\text{Total CO2 Eq. Emissions}_{i,t}, \text{Estimated Total CO2 Eq. Emissions}_{i,t})} \cdot 100\%. \tag{2.1}$$

It is imperative to note that in the above formulation, we are compelled to disregard cases where one or both of the variables assume a zero value or a negative value. We revisit these cases later in our analysis. We opt to ignore minor differences, defined as discrepancies smaller than 1% in value given that this might be the cause of rounding and such differences are not expected to affect the analysis in a significant way. After implementing this criterion, we identify 255 instances where the difference exceeds 1%.

Although it is not possible to consider each case separately, for illustrative purposes, we consider the 6 companies,encompassing 120 out of the 255 identified instances, where the difference surpasses 100%. A thorough analysis of these cases, involving a scrutiny of reported values based on published reports, is presented in the Appendix, Table C.3. Among these cases, one instance involves a holding company, with the corresponding subsidiary also present in our dataset. We opt to eliminate this pair of firms from the dataset. A more comprehensive analysis of similar pairs of companies and the rationale for their removal is discussed in Subsection 2.4.8.

Furthermore, within the six investigated instances, one firm is identified for which the emissions value cannot be further scrutinized due to a lack of available information in reports. This challenge becomes more prevalent in the subsequent phases of our analysis when manually checking reports. While many companies publish emissions data online, it is acknowledged that our resources are limited compared to those of LSEG Workspace, hindering our ability to extract emissions data from companies. Consequently, in cases where we suspect an error in emissions data but are unable to rectify it due to insufficient information, we opt to exclude the company from the dataset.

In addition, we identify two firms where the inconsistency between the two values arises from a wrong unit of measurement for one of the variables. Specifically, one of the variables is reported in kilograms or megatonnes instead of tonnes. These cases are rectified, and such discrepancies are notably discernible as reporting in the incorrect unit of measurement for one variable results in a difference of at least a factor of 1000 compared to the true value. Moreover, one firm exhibits inconsistency between the two variables due to a partial restatement of emissions. By comparing the reports of 2020 and 2022 for the concerned company, we discern that an update for scope 1 in fiscal year 2020 transpired in fiscal year 2022.

We highlight one case where the variables 'Total CO2 Equivalent Emissions', do not align with the report, while the variable 'Estimated Total CO2 Equivalent Emissions' does align. Notably, this inconsistency is not attributable to updated values or an incorrect measurement unit. In this particular case, while the root cause remains unidentified, we rectify the discrepancy.

Finally, we remove the other 135 instances for which the difference between the two emissions variables was found to be above 1%.

2.4.5 Comparing emissions in successive years, per firm

In the preceding analysis, we uncovered several inconsistencies in the data through a comprehensive comparison of the two emissions variables. We identified several root causes, including the partial restatement of values and the wrong unit of measurement.

Identifying errors stemming from partial restatement is challenging, particularly when the restatement results in minor value differences. In contrast, recognizing inconsistencies arising from incorrect units of measurement is more straightforward. Such errors constitute genuine data discrepancies and have the potential to significantly impact outcomes, causing reported emissions to be inflated or understated by a factor of at least 1000. Rectifying such instances, where possible, or excluding them otherwise, becomes paramount. While we addressed such cases in the previous section by comparing the two emission variables, scenarios where both variables were reported in the wrong unit for a firm and time point remained undetected.

Subsequently, we focus on the variable 'Total Estimated CO2 Equivalent Emissions' in the ongoing analysis, disregarding the variable 'Total CO2 Equivalent Emissions', as any substantial discrepancies between the two have been rectified or eliminated.[7] Therefore, moving forward, we refer to 'Total Estimated CO2 Equivalent Emissions' when discussing the emissions levels.

For the identification of other potential anomalies related to incorrect measurement units, we investigate abrupt changes in emissions levels between successive years for a given firm. Specifically, for each firm, we calculate the differences in emissions between adjacent time periods, provided both periods include emissions data. The difference between two consecutive time periods, at time t for firm i, is defined as:

$$\text{Difference}_{i,t} = \frac{|\text{Total emissions}_{i,t} - \text{Total emissions}_{i,t-1}|}{\min(\text{Total emissions}_{i,t}, \text{Total emissions}_{i,t-1})} \cdot 100\%. \tag{2.2}$$

It is crucial to highlight that, in instances where emissions are estimated in two consecutive years, confirming the accuracy of values becomes challenging. While such cases raise suspicion, rectification based on reports from the corresponding firms is unfeasible. Therefore, our attention is directed towards periods where either both emissions being compared were reported by the firm or one was reported, and the other was estimated.

[7] After eliminating disparities between the two emission variables, they are equal when emissions are reported, whereas only the variable 'Estimated Total CO2 Equivalent Emissions' is filled in when emissions are estimated. Thus, the 'Estimated Total CO2 Equivalent Emissions' variable contains all requisite information. We acknowledge that instances where one of the two variables equals zero have not been considered at this stage, and we will revisit these cases later in the analysis.

In our examination, we identify 690 instances where the difference in emissions between two consecutive years exceeds 1000%, with 55 cases surpassing $100,000\%$.[8] Each of the 55 instances, involving 51 different firms, is meticulously considered. For a detailed breakdown of each case, please refer to the Appendix, Table C.4.

Out of the 51 distinct firms where difference between successive years equals or exceeds $100,000\%$, we confirm suspicions of an error occurring in 11 cases. In all these instances, the root cause of the error is identified as the wrong unit of measurement for one of the two years. For 15 firms, where the required information on emissions is unavailable, we opt to remove these observations. Regarding other suspicious cases, in most instances, the discrepancy arises between years where emissions were estimated in one year and reported in the other. After scrutinizing the relevant reports, no inconsistencies in reported values are found for these firms. The substantial difference between consecutive years may result from either a genuine large change in emissions emitted by a firm or the estimated value deviating significantly from the true value in a particular year. However, it is crucial to note that vendor-estimated values cannot be rectified, and we do not intend to rectify or remove these cases. This underscores the inherent complexity of addressing errors in vendor-estimated values as an external party, highlighting one of the reasons why, in the subsequent phases of our analysis investigating the impact of emissions on returns, we differentiate between vendor-estimated and reported values.

Further analysis has been exclusively conducted for cases where the percentage difference in emissions levels for pairs of consecutive years is at least $100,000\%$. Although this threshold may be conservative, as differences smaller than $100,000\%$ could still indicate data errors, we set this criterion to focus on more extreme discrepancies. There are $3,676$ cases where the difference falls between 100% and $100,000\%$. Acknowledging the impracticality of manually examining each case, we recognize that such instances may genuinely occur. Therefore, we refrain from rectifying these cases.

2.4.6 Zero values in emissions

We shortly return to the comparison between the variables 'Total CO2 Equivalent Emissions' and 'Estimated Total CO2 Equivalent emissions'. In our earlier analyses of these two variables, our focus was exclusively on cases where both variables were non-zero. We further eliminate 63 cases where the reported value equals zero while the vendor-estimated value is non-zero.[9] Furthermore, we identify $2,506$ cases where the vendor-estimated emissions are equal to zero, while no emissions are reported. We do not remove these observations given that they are still estimated values.

[8] Instances where the difference is at least $100,000\%$ indicate pairs of consecutive observations differing by a factor of at least 1000, suggesting the possibility of one of the instances being reported by LSEG Workspace in the wrong measurement unit.

[9] One explanation for the zero emissions restatement by LSEG Workspace could be that it deemed these reported values implausible, leading to the estimation of these values.

2.4.7 Discrepancy between scope 1 plus scope 2 and total emissions

We return to the variable 'Total CO2 Equivalent Emissions' as our primary emissions variable, bearing in mind that, according to LSEG Workspace, this variable represents the aggregate of scope 1 and scope 2 emissions. To validate this assertion for all cases, we conduct a comparison between the individual scope 1 and scope 2 emissions and the aggregate as reported by LSEG Workspace.

Initially, we observe no instances where the aggregate emissions value is empty while scope 1 and scope 2 emissions are reported. Subsequently, we identify 60 cases where total emissions are equated to scope 1, while scope 2 is non-zero. In 24 of these cases, where scope 2 is smaller than 1, we attribute the equality to rounding effects. However, the remaining 36 instances where scope 2 exceeds the value 1, are considered as data errors. To rectify this, we adjust the aggregate value by setting it equal to the sum of scope 1 and scope 2 for each of these 60 cases. Additionally, there are 192 instances where the aggregate is equal to scope 2 while scope 1 is filled. In all these cases, scope 1 values are not below 1. Again, we rectify these 192 instances by adjusting the aggregate to equal the sum of scope 1 and scope 2 emissions.

Subsequently, we address instances where only scope 1 or scope 2 emissions are reported, and the total is equated to the reported scope. We identify $6,678$ instances where the total is equated to scope 1 while scope 2 is unreported and $5,479$ instances where the total is equated to scope 2 while scope 1 is unreported. While acknowledging that a company can have either direct (scope 1) emissions or indirect (scope 2) emissions without the other, we consider such cases as improbable in most instances. Consequently, we choose to remove these observations from the dataset.

Third, we compute the sum of scope 1 and scope 2 emissions for cases where scope 1 and scope 2 emissions were both reported. We then compare the difference between the computed sum of scope 1 and scope 2 and the aggregate emissions as reported by LSEG Workspace. Specifically, for firm i at time t we calculate the difference as,[10]

$$\text{Difference}_{i,t} = \frac{|(\text{Scope 1} + \text{Scope 2})_{i,t} - \text{Aggregate emissions}_{i,t}|}{\text{Aggregate emissions}_{i,t}} \cdot 100\%. \tag{2.3}$$

We identify $10,972$ cases where the aforementioned difference exceeds 1%. Smaller deviations are ignored, as they are likely attributable to rounding errors, as noted by Bajic et al. (2023). Conversely, larger discrepancies, as discussed by the aforementioned authors, often stem from restatements of emission values in new reports published by firms. Typically, these restatements apply to the aggregate value and not separately to scope 1 and scope 2 emissions. Other potential reasons for significant deviations include differences in measurement units, errors resulting from only partial restatement of emission values, variations in reporting coverage for the aggregate compared to scope 1 and scope 2, inconsistencies in values reported by the companies themselves, and incorrect updates by LSEG Workspace. Analyzing the root cause for each instance and potentially rectifying these data issues would necessitate a meticulous examination of each case individually. Given the substantial number of instances, conducting such an analysis is impractical. Consequently, we choose to remove these observations from the dataset.

[10]We acknowledge there are cases where the aggregate emissions are equal to zero, as discussed previously. However, in all these cases, scope 1 and scope 2 are not filled, rendering the difference calculation inapplicable in any case.

Last, it is essential to note that we did not remove cases where total emissions include a valid value while both scope 1 and scope 2 are not reported. In most cases, this scenario arises when total emissions are estimated by the data vendor, and scope 1 and scope 2 are not separately reported. However, there are cases where total emissions were reported by the firm according to the estimation method, but scope 1 and scope 2 values are not available in the data. As emphasized by Bajic et al. (2023), many companies do not report scope 1 and scope 2 emissions separately. We choose to retain these observations in the dataset, assuming that they encompass both scope 1 and scope 2 emissions.

2.4.8 Dual listings and subsidiaries

The discussion on emissions data concludes with an examination of dual-listed companies and subsidiaries, following the discussion by Bajic et al. (2023). Challenges may arise when the data encompasses both the holding company and its subsidiary or involves dual-listed companies. While the presence of both the holding company and the subsidiary is not inherently problematic, discrepancies in other accounting aspects, such as reported sizes or financial figures, can lead to divergent carbon footprints, particularly when scaling emissions with a measure of company size. In situations involving dual-listed companies, the presence of identical emission values alongside accounting revenue and other financial figures may result in duplicates. This concern, as emphasized by Bajic et al. (2023), becomes particularly important for large firms or when making regional or industrial comparisons, especially if the company operates in a country on industry underrepresented in the sample.

To mitigate these concerns, we identify potential parent/subsidiary pairs and dual-listed companies in our dataset by searching for equivalent emissions values in a given month. Specifically, we mark a pair of companies as potentially dual-listed or as a parent-subsidiary pair if they share the same emissions values for at least 24 months. We identify 82 such pairs. Determining if a pair truly consists of dual-listed companies or if one is the holding company and the other is the subsidiary would require a more thorough analysis, including identifying the structure of each company separately. However, such a detailed analysis is beyond the scope of this study. In Table C.5 we present three cases involving both parent/subsidiary pairs and dual-listed companies. Due to the potential for double counting and misleading carbon footprints, we choose to remove the 82 identified suspicious pairs of companies from the dataset.

2.4.9 Winsorization

We refrain from winsorizing the emissions variable. As noted by Bajic et al. (2023), important insights, especially regarding sectoral variations, may reside in the tails of the distribution data. Winsorizing emissions data could inadvertently result in overlooking significant polluters, potentially excluding valuable information. Furthermore, in the data cleaning process, we have already compared firms' emissions between consecutive years to identify values that may have resulted in faulty outliers. On the other hand, we do winsorize financial and accounting data, according to Table C.6, in accordance with winsorization applied by BK.[11]

[11] We note that winsorizing control variables is generally common in the literature. We do not elaborate on this part further given that Aswani et al. (2023) and BK find that their inferences ar robust to different winsorization levels. As we focus on other factors that might potentially affect inference, we keep the winsorization as close as possible to other papers in order to compare our results to other papers.

Chapter 3

Descriptive properties

This section highlights several noteworthy points in our data. Amongst others, we present summary statistics of our global sample, and subsequently discuss the distribution of firms in our sample across countries and over time. Furthermore, to bridge the gap to our analysis, we discuss the effect of estimating emissions on the level of emissions and consider the relation between emissions and firm characteristics.

3.1 Summary statistics

In Panel A of Table C.7, we provide comprehensive summary statistics for global emissions data. Notably, there is significant variation in both total emissions and emissions intensity, primarily due to differences between firms rather than within individual firms over time. However, within-firm variation over time remains substantial, highlighting the dynamic nature of emissions data.

Our dataset, after cleaning emissions data, comprises $831,314$ observations from $10,516$ firms across 90 countries[1]. This wide coverage spans diverse countries and exceeds the number of firms included in most other papers, except for the paper by BK2, which includes $14,400$ firms across 77 countries. Our sample encompasses slightly more countries than BK2, reflecting the inclusion of more recent years when emissions reporting expanded to developing countries.

3.2 The effect of data cleaning on emissions data

In Section 2.4, we extensively cleaned the emissions data. In Panel B of Table C.7, we compare descriptive statistics of emissions data before and after cleaning. On average, emissions levels decreased by nearly 15% after cleaning. In addition to the decrease in average emissions, we also notice a substantial decrease in the overall standard deviation of emissions levels, which can be mainly attributed to a reduction in within-firm variation in emissions, for example due to comparing emissions in successive years. In many of these cases, we either corrected the data or removed observations that could not be rectified. Negative emissions values, stemming from seven observations for one firm, were also removed during cleaning.

[1] These number slightly differ from the Figures presented in Figure B.1 in the last step. This is due to the removal of one company from Thailand and one company from Indonesia as a result of extremely high emissions intensity. Given the relatively small size of the number of observations in these countries, we deemed it important to remove these two observations.

Finally, when considering emissions intensity, the effects of the data cleaning process are less pronounced in terms of average emissions intensity. One reason for this may be that some of the largest values of emissions intensity were not removed from the dataset since the corresponding emissions levels were not detected as suspicious cases during the data cleaning process.

However, there is a noticeable decrease in the standard deviation of emissions intensity, primarily due to reduced within-firm variation after cleaning. This suggests that large fluctuations in emissions levels between consecutive years, flagged as errors, also imply significant variations in emissions intensity. If these fluctuations were due to extreme growth rather than errors, emissions intensity should not have increased drastically, and such cases should not have been flagged.

3.3 Country comparison

We observe an uneven distribution of our sample across countries and time, as evident from Table C.8 and Figure B.2. Table C.8 presents a breakdown of the distribution of firms within our sample at the country level, along with relevant statistics. The United States notably dominates the sample, accounting for almost one-third of the total observations. Following, China constitutes approximately 7.4% of observations, with Japan as the third-largest contributor at about 6.4%.

Notably, the top three countries exhibiting the highest average log emissions per firm are Russia, Ukraine, and Monaco. However, we acknowledge the limited number of observations for Ukraine and Monaco, cautioning that considering these countries in isolation may not provide a representative portrayal. Among the major nations, Japan, China, and Germany showcase relatively high average emissions, while Australia and Sweden demonstrate more favorable standings when evaluating log emissions. Intriguingly, a different perspective emerges when examining emissions intensity. Canada, Monaco, and Russia top the list as the most intense countries, with average emissions intensities of 2.91, 2.51, and 1.65 tonnes of CO_2e per USD dollar in revenues, respectively.

This comparison highlights the substantial influence of scaling emissions. For instance, while The Netherlands may initially appear relatively pollutant when assessing emissions levels alone and ranking 54st for the lowest average emissions levels, its position improves to 22nd when considering emissions intensity scaled by revenues.

3.4 Firm coverage over time

We analyze the temporal evolution of firms' coverage in our sample, focusing on emissions data. Figure B.2 in the Appendix shows the count of companies with emissions data, both reported and vendor-estimated, and scope 1 and scope 2 emissions separately. Notably, the count of companies with returns data increases over time, reflecting the growing number of companies contributing to our dataset as it is based on those active when the lists were created by LSEG Workspace in 2022. The graph also displays a close alignment between the number of firms reporting scope 1 and scope 2 emissions for most periods. The small discrepancies observed suggest that companies reporting on one scope typically report on the other, indicating a strong correlation in disclosure practices.

In addition, as many firms disclose emissions at year-end, the monthly emission count remains relatively stable throughout the year. However, periodic changes occur due to variations in reporting intervals among some firms.

Moreover, a notable gap exists between companies reporting aggregate greenhouse emissions and those disclosing granular data separated into scope 1 and scope 2 emissions. While the number of companies reporting aggregate emissions consistently surpasses those reporting scope 1 and scope 2 separately, recent years show a narrowing gap, suggesting an increasing trend of comprehensive reporting.

Last, the graph highlights a substantial increase in available emissions data over the sample period, reflecting both an expansion in the number of firms and an improved trend in reporting behavior. Notably, the percentage of companies reporting total emissions rose from 36.3% in February 2007 to approximately 62.1% in February 2021. This trend aligns with the global increase in climate-related financial disclosures and emphasizes the growing importance of emissions data in recent years (TCFD, 2022). However, a decrease in reporting is observed in 2022 and January 2023, possibly indicating a lag in emissions disclosure by some companies as well as potential effects from the recent events.

3.5 Reported versus vendor-estimated emissions

As highlighted in Appendix Section A.1, the inclusion of disclosed and vendor-estimated emissions data, and its methodological implications, have sparked significant debate. This section explores observed differences between vendor-estimated and reported emissions.

In Panel B of Table C.7, we present separate figures for vendor-estimated versus reported emissions. Notably, average emissions levels are relatively low for the 'Median' method compared to other estimation methods and reported figures. Differences are more pronounced in emissions levels than in log emissions levels, with 'Energy' method emissions levels averaging over 17 times those from the 'Median' method.

Reported emissions exhibit relatively high average levels, which could stem from several factors. Systematic disparities between firms reporting emissions and those relying on estimates may exist, suggesting stricter regulations for emissions disclosure among more polluting firms. However, this explanation falls short in elucidating the higher average emissions based on the 'Energy' method. The higher average emissions from the 'Energy' method may imply that emissions from non-disclosing highly polluting firms are often estimated using this method. Alternatively, and more in line with existing literature on estimation methods by data vendors, the observed discrepancies might suggest that estimation methods, in general, do not provide accurate estimations of emissions levels (e.g., Aswani et al., 2023; Busch et al., 2020). One potentially supportive observation is the relatively high number of emissions estimated by the final source method, the 'Median' method. When none of the preferred methods provide an estimate, the 'Median' method is used as a last resort, implying potential lower accuracy compared to other methods.[2]

A second striking point arises from the within variation for reported emissions. Although we recognize the variability of emissions over time for a given firm, the within variation for reported

[2] This is speculative, based on LSEG Workspace's use of the 'Median' method as a last resort. Further research into individual methods is necessary to explore potential systematic differences.

emissions is relatively high compared to emissions that are vendor-estimated. This result raises questions on the reliability of reported emissions.

To formally examine the observed differences between reported emissions and the vendor-estimated emissions, accounting for potential differences in estimation methods, we employ a regression model, expressing emissions levels as a function of the estimation method and relevant firm characteristics in the US sample and the global sample:

$$\text{Emissions}_{i,t} = \beta_0 + \beta_1 \text{CO2}_{i,t} + \beta_2 \text{Energy}_{i,t} + \beta_3 \text{Median}_{i,t} + \delta \text{Controls}_{i,t} + \lambda_{industry} + \gamma_t + \epsilon_{i,t}. \quad (3.1)$$

Here, $\text{Emissions}_{i,t}$ denotes a generic measure, encompassing either log emissions levels or emissions intensity. Log emissions levels are measured as $\log(1 + CO_2\text{e emissions})$ and emissions intensity is measured as CO_2e emissions scaled by revenues. $\text{CO2}_{i,t}$, $\text{Energy}_{i,t}$, and $\text{Median}_{i,t}$ are indicators denoting whether the estimation method for a given firm i at time t is 'CO2', 'Energy', or 'Median', respectively. The control group comprises emissions that are reported, and we account for firm characteristics through $\text{Controls}_{i,t}$. These characteristics, also considered in the subsequent analysis of the relationship between returns and emissions variables, aim to control for inherent differences across firms. It is important to note that we add log revenues to the control variables in this regression.[3] Additionally, we include industry-fixed effects, $\lambda_{\text{industry}}$, and month-year fixed effects, γ_t, to capture industry-specific and time-specific factors. We perform similar analyses for the global sample, but then also include country-fixed effects.

The regression results based on log emissions levels are presented in Column (1) of Table C.9 for the US sample and in Column (2) of the same table based on the global sample. Similarly, results based on emissions intensity for the two samples are included in Columns (3) and (4). Controlling for firm characteristics, industry-fixed effects, and time-fixed effects, we observe no statistically significant differences between reported and vendor-estimated emissions levels for the US, while for the global sample, emissions based on the 'CO2' method are on average lower than reported emissions levels, when controlling for firm-specific variables and fixed effects. Based on columns (3) and (4) we find that emissions intensities based on the 'Median' method are on average lower than emissions intensities based on reported emissions for the US, but globally we find no effect of the estimation methods on emissions intensities. These findings indicate to some extent the existence of a systematic difference between reported and vendor-estimated emissions, not solely attributable to firm characteristics.

This raises concerns akin to those articulated by Aswani et al. (2023) regarding the inclusion of both reported and vendor-estimated emissions when exploring the potential relationship between a firm's returns and emissions data. However, the difference between reported emissions and vendor-estimated emissions seems to depend on specific estimation methods, on the geographic region included and whether or not emissions are scaled by revenues.

The findings in Columns (1) and (2), based on unscaled emissions, of Table C.9 underscore another observation. Specifically, emissions levels exhibit a substantial dependency on firm characteristics. Further elaboration on this aspect follows.

[3] Revenues are not added in columns (3) and (4) of Table C.9 given that the dependent variable is scaled by revenues in these regressions.

3.6 Financial information captured by emissions data

Finally, we consider the extent to which emissions depend on companies' accounting data, as found by Zhang (2024). As suggested by the regression results in Table C.9, emissions levels are dependent on a firm's characteristics. We find that over 90% of the variation in log emissions levels can be explained by variation in firm's characteristics and by time-fixed effects and industry-fixed effects for the US sample. For the global sample, this value equals to 81.6%

Zhang (2024) notes that the generalized methodological approach to calculate emissions suggests that calculated emissions are highly dependent on firms' sales. In particular, the most common approach used by firms to calculate[4] GHG emissions is through the application of documented emissions (WBCSD and WRI, 2004). That is,

$$\text{Emissions levels} = \text{Activity data} \times \text{Emissions factor}. \tag{3.2}$$

In this equation, activity data are data on a firm's activity resulting in emissions, such as electricity use and gasoline use.[5] The emissions factor is a coefficient that describes the rate at which a given activity releases greenhouse gases.[6] A final step in the calculation is to express emissions into CO_2 equivalents, which requires an additional calculation.[7]

The above equation suggests that emissions are highly dependent on a firm's accounting data, in particular sales or revenues, in case firms calculate emissions using this formula. Similar as to Zhang (2024), we test to what extent this holds true in our dataset. In this respect, we analyze the information on revenues contained in emissions data by regressing log revenues on log total emissions, for firm i at time t,

$$\text{Log(revenues)}_{i,t} = \beta \text{Log}(1 + \text{total emissions levels})_{i,t} + \epsilon_{i,t}. \tag{3.3}$$

In the above, $\epsilon_{i,t}$ is the idiosyncratic error term. The results are presented in Columns (1)-(3) of Table C.10 in the Appendix. Panel A presents the results for the US sample, while Panel B includes the results for the global sample. In Columns (4)-(6), we additionally add time-fixed effects and industry-fixed effects. For the global sample, we also include country-fixed effects in those three columns. We furthermore distinguish between a sample incorporating reported and vendor-estimated emissions, reported emissions only and vendor-estimated emissions only.

For each of the six specifications, and both for the US and the global sample, we find that emissions grow with sales. Considering the US sample in Panel A, revenues explain over 64% of the variation in emissions when considering both reported and vendor-estimated emissions. Notably, this number reduces to approximately 33% when considering reported emissions and equals about 63% when focusing on vendor-estimated emissions. If we add time-fixed effects

[4] To be precise, the firms reporting emissions based on the formula provided, are actually estimating their emissions. However, in order not to confuse this approach with the approach taken by data vendors to estimate firms' emissions, we still refer to these emissions as 'reported' emissions.

[5] The type of activity data depends on the scope of emissions being calculated.

[6] The sources of emissions factors include the GHG protocol, IPCC and several government agencies. Typically, these factors are merely averages derived from all available data of satisfactory quality and are commonly presumed to be indicative of long-term averages applicable to all facilities within the source category. For more information, see `https://www.epa.gov/air-emissions-factors-and-quantification`.

[7] Different greenhouse gases are expressed into CO_2 equivalents using The Global Warming Potential (GWP) approach, since different gases contribute differently to global warming. For more information, see `https://www.epa.gov/ghgemissions/understanding-global-warming-potentials`.

and industry-fixed effects, these numbers are raised to 90.2% for the full sample, 79.2% for reported emissions and 91.1% for vendor-estimated emissions. In Column (1) of Table C.9, we found that firm characteristics, including revenues, can explain 90.6% of the heterogeneity in firm's reported and vendor-estimated emissions. That is, adding additional firm characteristics, besides revenues and fixed effects, increases the explanatory power by only 0.4% points implying that revenues, time and industry can account for a relatively large part in the heterogeneity of emissions already. Similar results are found for the US, with the latter difference being less than 0.1% points.

As discussed in Subsection A.1.5 in the Appendix, carbon emissions are often reported by firms with considerable lags. In case one would not sufficiently lag emissions data while accounting data is lagged, the above analysis would suggest a forward-looking bias. Furthermore, the above analysis suggests that a potential link between log level emissions and stock returns may be the result of a link between a firm's size and stock returns.

We note that Table 4 of BK2 can explain about 54% of variation in global scope 1 and scope 2 emissions when focusing on reported and vendor-estimated emissions, including several firm characteristics and fixed effects, excluding log revenues. In particular, we can explain about the same amount of variation in emissions in our sample by solely focusing on log revenues, excluding fixed effects.

The above suggests a high dependency of emissions levels on firm characteristics. An important question in this respect is if scaling emissions by a firm's size can diminish this dependency. We therefore return to Columns (3) and (4) of Table C.9. Several points emerge from this Table. First, when considering emissions intensity instead of log total emissions levels, the coefficients of most of the firm characteristics are insignificant at the 1% level and the variation in log emissions levels explained by firm characteristics and industry effects reduces to 0.38% in the US and 0.031% globally. That is, scaling emissions by revenues purges out most of the company information contained in emissions data and this effect is stronger in the global sample compared to the US, an argument that has been used against the use of emissions levels and in favor of using emissions intensity, see Subsection A.1.4 in the Appendix.

The results from the above analysis highlight a second point. In particular, they suggest high correlation between emissions levels and variables that capture a firm's size. Considering this in more detail, we find a correlation between log emissions levels and log revenues equal to 0.51 in the US and 0.46 in the global sample, when including both vendor-estimated and reported emissions. We note that BK and Aswani et al. (2023) account for firms size as a control variable through the log of market capitalization in their panel regressions. Given that we expect market capitalization to move along with a firm's revenues, and revenues are highly correlated with emissions, we therefore also consider the correlation between log emissions levels and the log of market capitalization. The correlation between log emissions levels and the log of market capitalization equals 0.64 in the US and 0.54 in the global sample. Once emissions are scaled by log revenues, both correlations are close to zero, for the US as well as for the global sample. Given the high correlation between log emissions levels and log market capitalization, including log emissions levels and log market capitalization as independent variables in panel regressions may induce a risk of strong multicollinearity.

Chapter 4

Methods

In this chapter, we discuss the models utilized to assess the impact of a firm's environmental performance on stock returns. Specifically, we introduce panel regressions and portfolio analyses, and finally replicate a methodology proposed by Bauer et al. (2022). Additionally, we discuss the temporal lagging of emissions and accounting data, alongside the introduction of a metric aimed at gauging shocks to climate change concerns.

4.1 Panel regression analysis

We initiate our analysis with a baseline regression model. In particular, we estimate the following panel regression,

$$\text{Ret}_{i,t} = \alpha + \beta\text{Emissions}_{i,t} + \delta\text{Controls}_{i,t} + \mu_t + \epsilon_{i,t}. \tag{4.1}$$

This regression operates at the firm-month level, where $\text{Ret}_{i,t}$ measures the stock returns of firm i at time t and $\text{Emissions}_{i,t}$ is used as a generic term specifying either log emissions levels or emissions intensity[1] of firm i at time t. The term μ_t captures time-fixed effects at the month-year level. $\text{Controls}_{i,t}$ constitutes a vector of accounting variables, encompassing firm characteristics demonstrated to be associated with stock returns, primarily following the framework of BK. We refer to Table C.6 in the Appendix for an overview of the included control variables. Our primary analysis spans from February 2007 until January 2023.

Across all model specifications, standard errors are double-clustered at the firm and month levels. This approach accommodates any cross-firm correlation in the residuals and acknowledges the temporal correlation of certain variables, such as emissions, often measured annually. Moreover, in the majority of our specifications, we introduce industry-fixed effects to capture within-industry variation across firms. Explicitly denoting industry-fixed effects as $\lambda_{industry}$, the refined model specification becomes,

$$\text{Ret}_{i,t} = \alpha + \beta\text{Emissions}_{i,t} + \delta\text{Controls}_{i,t} + \mu_t + \lambda_{\text{industry}} + \epsilon_{i,t}. \tag{4.2}$$

Industry classifications are based on the Refinitiv Business Classification (TRBC) scheme, we refer to Table C.6 for a detailed explanation. Lastly, when conducting analyses for our global sample, country-fixed effects are incorporated. Again, Table C.6 details the country information utilized in our study.

[1] In all our specifications, emissions intensity is quantified as CO_2e emissions scaled by revenues.

4.2 Portfolio analysis

Several referenced studies, including Bauer et al. (2022) and Pastor et al. (2022), emphasize the significance of portfolio analysis. As highlighted by Bauer et al. (2022), studies suggesting a carbon premium often lean towards panel regressions, while evidence favoring green outperformance is frequently associated with portfolio analyses. To gauge the potential impact of the model on the results concerning the link between a firm's environmental performance and its stock returns, we incorporate portfolio analysis. This requires a minimal expansion of our data, incorporating control variables, particularly factors from the Fama-French 5-factor model and the momentum factor (Carhart, 2012; Fama and French, 2015).

On a monthly basis, we categorize stocks into quintiles based on either their emissions levels or computed emissions intensity. Subsequently, we calculate the spread return on our Green-Minus-Brown (GMB) portfolio using value-weighted returns. This involves establishing a long position in the quintile of value-weighted stocks with the highest levels of greenness and a short position in the quintile of value-weighted stocks with the lowest levels of greenness. To accommodate recognized risk factors potentially influencing stock returns, we conduct the following regression:

$$\mathrm{Ret}_t^{\mathrm{GMB}} = \alpha + \beta_1 \mathrm{RMRF}_t + \beta_2 \mathrm{SMB}_t + \beta_3 \mathrm{HML}_t + \beta_4 \mathrm{RMW}_t + \beta_5 CMW_t + \beta_6 \mathrm{UMD}_t + \epsilon_t. \quad (4.3)$$

Specifically, we regress the GMB portfolio's monthly return on the monthly factors integrated into the Fama-French 5-factor model and the momentum factor.[2] The Fama-French 5 (FF5) factors comprise the excess market return at time t (RMRF), the high-minus-low factor (HML), the small-minus-big factor (SMB), the robust-minus-weak factor (RMW), and the conservative-minus-aggressive factor (CMW) (Fama and French, 2015). Additionally, we include the up-minus-down (UMD) factor, also known as the momentum factor, proposed by Carhart (2012). The intercept, denoted as α, serves as the coefficient of interest, capturing the return differential between green and brown assets, based on either emissions levels or emissions intensity, unexplained by systematic risk factors.

[2] The Fama-French 5-factors and Momentum factors are taken from the data library available on the official Kenneth R. French website, see `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

4.3 Replicating the method of Bauer et al. (2022)

While panel regressions and portfolio analyses exhibit fundamental differences, Bauer et al. (2022) propose a methodology aimed at partially reconciling the distinctions between these two approaches. In regression analyses, a continuous measure of a firm's emissions serves as the independent variable, whereas in portfolio analyses, emissions data indirectly sorts stocks into quintiles, resulting in a spread utilized as the dependent variable. To bridge these methodological gaps, Bauer et al. (2022) devise an emissions indicator variable, akin to the sorting process in portfolio analyses, which is then incorporated into panel regression analysis.

We establish a metric for a firm's environmental performance mirroring the approach of Bauer et al. (2022), for two primary reasons. First, to facilitate comparison between the results of panel regression and portfolio analyses. Second, as the only existing study examining the correlation between a firm's emissions and stock returns utilizing data from LSEG Workspace is conducted by Bauer et al. (2022)[3], we deem it imperative to juxtapose our findings with theirs and, to a certain extent, replicate their methodology. Bauer et al. (2022) introduce an indicator variable, assigning a value of +1 to stocks falling within the highest quintile of emissions or intensity, and −1 to those in the lowest quintile. Stocks in intermediate quintiles are assigned a value of 0. As highlighted by Bauer et al. (2022), regressing this indicator variable on monthly stock returns in a panel regression is akin to computing the mean return of an equally weighted GMB spread.[4] To further incorporate firm-specific characteristics, time, and industry-fixed effects, we regress stock returns on the constructed indicator variable and control variables, following our baseline specification. To additionally account for firm-specific characteristics and time and industry-fixed effects, we regress stock returns on the constructed indicator variable and control variables as in our panel regressions introduced earlier,

$$\text{Ret}_{i,t} = \alpha + \beta \text{Emissions indicator}_{i,t} + \delta \text{Controls}_{i,t} + \mu_t + \lambda_{industry} + \epsilon_{i,t}. \tag{4.4}$$

In the above, we define, opposite to Bauer et al. (2022)[5], the indicator on emissions to take value +1 if firm i belongs to the greenest fifth of firms based on either emissions levels or emissions intensity at time t, −1 if it belongs to the brownest fifth and 0 otherwise. Since we perform this analysis for the US only, we do not account for country-specific effects.

While this specification closely resembles the one in Equation 4.2, we now use the indicator variable on emissions as a gauge of a firm's environmental performance. Unlike annual re-classification of stocks applied by Bauer et al. (2022), we undertake this re-classification monthly. This monthly sorting is essential, considering that emissions may be reported at different intervals within the same year for different firms.

[3] We acknowledge that Trinks et al. (2022) also utilize LSEG Workspace's emissions data in their analysis, albeit considering cost of equity as their dependent variable instead of stock returns.

[4] The disparity between computing the mean return of the equally weighted GMB spread and the estimated coefficient on the indicator variable stems from differing weighting procedures. While the mean return of the equally weighted GMB spread assigns equal weight to each month, regardless of the number of observations, panel regressions give more weight to months with greater observations. Hence, the estimated coefficient on the indicator variable in the regression of stock returns may not precisely correspond to the average equal-weighted GMB spread.

[5] We choose to model the indicator taking value +1 for the greenest fifth of stocks given that we previously defined a Green-Minus-Brown portfolio, going long in the greenest quintile of stocks and shorting on the brownest fifth of stocks. Bauer et al. (2022) defined exactly an opposite portfolio, that goes short in green stocks and long in brown stocks.

4.4 Lagging emissions and accounting data

Emissions data often experience a delay in reporting, a topic extensively discussed in the literature, see Subsection A.1.5 in the Appendix. In some of our analyses, we therefore align emissions data closer to investor information available during investment decisions. However, the lack of individual company-specific lag information in LSEG Workspace poses a challenge. Consequently, when implementing lagged emissions data, we resort to lag information from previous literature. Median emissions lags, as noted by Zhang (2024), extend to 10 to 12 months post-fiscal year-end for both US and international companies. Notably, Bauer et al. (2022) underline that emissions reporting commonly occurs six months after the conclusion of the fiscal year, akin to the reporting timelines of accounting and financial data. We note that these provided lags are typically averages or median values and may not precisely match the reporting timelines of specific companies. Hence, our analysis investigates different lag durations for emissions data, with a specific focus on intervals documented in existing literature.

Moreover, accounting data typically become available six months subsequent to the conclusion of the fiscal year. Consequently, in certain segments of our analysis, a lag of 6 months is incorporated for accounting data to align with reporting timelines. Notably, as observed by Zhang (2024), emissions data exhibit a lag behind accounting data. To address possible forward-looking biases caused by lagging emissions data less than accounting data, while still incorporating financial information, we ensure that the lags applied to emissions data are at least as extensive as those applied to accounting data. It is worth noting that accounting data are solely employed in panel regressions, making this adjustment irrelevant for portfolio analyses.

4.5 Accounting for shocks via climate news

Most empirical studies investigating the link between a firm's environmental performance and its financial outcomes focus on realized returns, analyzing ex post data. The PST model, proposed by Pastor et al. (2019), proposes that green stocks may demonstrate higher realized returns owing to unforeseen shifts in demand towards environmentally friendly assets, while encountering lower expected returns. Consequently, the outperformance of green stocks compared to brown stocks in such scenarios may stem from an unexpected upsurge in environmental concerns, potentially misleading future predictions.

To assess the influence of investor climate concerns on our results, we integrate a measure of climate change concern into our established models. Specifically, we utilize shocks in the Media Climate Change Concerns index (MCCC) introduced by Ardia et al. (2022). This daily index spans from 2003 until June 2018 and is derived from climate change-related news articles from major US newspapers and newswires.[6] It quantifies concerns by analyzing the volume, risk focus, and negativity of these articles. The MCCC score is aggregated at the day level per newspaper, and the final score is computed as the square root of the aggregated values across different newspapers, adjusted for heterogeneity. This provides the MCCC as a daily proxy for climate change concern, with higher values indicating heightened concern.

[6]Sources include the New York Times, Washington Post, Los Angeles Times, Wall Street Journal, Houston Chronicle, Chicago Tribune, Arizona Republic, USA Today, New York Daily News, New York Post, Associated Press Newswires, and Reuters News.

However, the PST model concentrates on unexpected fluctuations in climate change concerns. Acknowledging that some climate news may be anticipated, a mere difference in the MCCC measure between two periods may not fully capture unforeseen shifts in climate change concerns. Hence, as a final step, Ardia et al. (2022) derive the unexpected shock component of the MCCC index as the prediction error of a first-order autoregressive model, incorporating financial-market, energy-related, and macroeconomic variables. This unanticipated measure, denoted as the Unexpected Media Climate Change Concerns (UMC), is utilized in our analysis.[7] Although the UMC measure is available on a daily basis from December 2008 until June 2018, we employ the monthly version due to the monthly focus of our analysis.[8] Given that a high (low) MCCC value indicates elevated (reduced) climate change concerns, a positive (negative) UMC value suggests heightened (diminished) climate change concerns.

We integrate the UMC factor into both our panel regression and portfolio analyses. Specifically, we extend the panel regressions equation given by Equation 4.2, by incorporating the UMC factor interacted with the emission measure in Equation 4.2, resulting in:

$$\text{Ret}_{i,t} = \alpha + \beta \text{Emissions}_{i,t} + \rho \text{Emissions}_{i,t} \cdot \text{UMC}_t + \delta \text{Controls}_{i,t} + \lambda_{\text{industry}} + \mu_t + \epsilon_{i,t}. \quad (4.5)$$

Here, $\text{Emissions}_{i,t}$ remains a versatile measure, representing either emissions levels or intensity. The term $\text{Emissions}_{i,t} \cdot \text{UMC}_t$ evaluates firms' exposure to the UMC factor, where ρ gauges a firm's exposure depending on its levels of emissions. According to the PST concept, which suggests green firms may outperform brown firms amid increasing climate change concerns, an anticipated negative coefficient (ρ) is expected on the interaction term, capturing changing preferences. The inclusion of $\text{Emissione}_{i,t}$ separately acknowledges that the UMC factor may only reflect heightened concerns, not capturing other factors such as regulatory transition risk.

Subsequently, in our portfolio analysis, inspired by Pastor et al. (2022), we incorporate the UMC factor as an additional regressor. We amend Equation 4.3 to include the UMC factor at time t, yielding:

$$\text{Ret}_t^{\text{GMB}} = \alpha + \beta_1 \text{RMRF}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{RMW}_t + \beta_5 CMW_t + \beta_6 \text{UMD}_t + \psi \text{UMC}_t + \epsilon_t. \quad (4.6)$$

Here, $\text{Ret}_t^{\text{GMB}}$ denotes the return of the GMB portfolio, with stocks sorted monthly based on emissions levels or intensity. As per the PST model, a larger UMC factor, indicative of heightened climate change concerns, suggests worse prospects for brown stocks compared to green ones, predicting a positive value of ψ.

[7]For a detailed explanation of the MCCC measure calculation and the corresponding UMC values, refer to the original paper by Ardia et al. (2022).

[8]The data on the MCCC values and the corresponding UMC values computed by Ardia et al. (2022) are publicly accessible as part of the Supplementary Materials from `https://pubsonline.informs.org/doi/10.1287/mnsc.2022.4636`.

Chapter 5

Results

In this chapter, we present our findings regarding the relationship between a firm's environmental performance and its stock returns. We highlight the importance of methodological decisions on our research outcomes, initially examining evidence from the US before extending our analysis globally. Our investigation unfolds through several stages. Firstly, we conduct firm-level panel regression analyses to study the correlation between emissions measures and returns while controlling for firm-specific characteristics. Subsequently, we delve into portfolio analyses, assessing a Green-minus-Brown spread and accounting for systematic risk factors. In both analyses, we implement various adjustments to test the robustness of results to methodological choices.

5.1 US evidence

5.1.1 Panel regression analysis

Given the predominant focus on the US within existing literature, particularly as a benchmark for analysis, our initial investigation centers on US-based firms. The temporal scope of all our analyses spans from February 2007 to January 2023, unless specified otherwise. All results are interpreted based on a 5% significance level.

We initiate our investigation by regressing returns on two distinct emissions measures, emissions levels, and emissions intensity, while controlling for firm-specific characteristics as outlined in Equations 4.1 and 4.2, and present the results in Panel A and B of Table 5.1, respectively. In each of the panels included in the Table, we only report the estimated coefficient on the emissions variable, being either log emissions levels or emissions intensity. We present the estimated coefficients of the remaining control variables for each panel in Table D.1 of the Online Appendix, corresponding to the results outlined in Table 5.1.

In Panel A, without industry-fixed effects, we find some evidence in favor of a negative carbon premium, particularly when considering firms with vendor-estimated emissions and focusing on emissions levels. Conversely, when employing emissions intensity, a positive carbon premium is observed, particularly with reported emissions, albeit not statistically significant for vendor-estimated emissions at the 5% significance level. However, one would expect differences between industries and including industry-fixed effects aligns with previous literature, given the prevailing analytical focus on studies incorporating industry-fixed effects, including BK, Zhang (2024), Bauer et al. (2022) and Aswani et al. (2023). The inclusion of industry-fixed effects reveals notable variations in results compared to Panel A, as outlined in Panel B.

With industry-fixed effects incorporated, as outlined in Equation 4.2, we find a statistically significant and negative coefficient on log emissions, driven by vendor-estimated emissions, which diminishes upon scaling emissions by revenues. Notably, distinguishing between reported and vendor-estimated emissions yields an economically meaningful difference in green outperformance. For instance, Panel B suggests that a US-based firm with vendor-estimated emissions equal to the 20th percentile value of log emissions, on average enjoys monthly stock returns that are 2.18% points higher compared to a similar firm with vendor-estimated log emissions equal to the 80th percentile. Conversely, the results suggest no statistically distinguishable effect for firms with reported emissions.[1,2]

Scaling emissions levels by revenues, we find no statistically significant relationship between emissions intensity and stock returns in Panel B. This observation concurs with the proposition that the link between stock returns and emissions data may primarily stem from the relationship between a firm's size and its stock returns, or potentially from concerns regarding multicollinearity. The latter concern arises due to the inclusion of emissions levels and the logarithm of market capitalization as regressors within a single regression model, consistent with the conclusions drawn by Aswani et al. (2023).

Thus far, we find no evidence of a positive carbon premium, in contrast to BK. Acknowledging that our sample encompasses a broader timeframe than that of BK and Aswani et al. (2023), we present results for the period from February 2007 to the end of 2018, in Panel C.[3] Despite differences in the magnitudes of coefficients, our overarching conclusions remain unchanged. Furthermore, these results indicate that the evidence supporting a negative carbon premium, is not dependent on the inclusion of the most recent years in our sample. These years have witnessed several extreme events that, according to the PST framework, could have influenced a shift towards greener firms.

Continuing our investigation, we extend our analysis to explore potential lags in emissions and accounting data, which may influence an investor's knowledge set. In Panel D, we lag emissions data and accounting data by 6 months. Notably, green outperformance based on vendor-estimated log emissions levels persists, but the effect is not strong enough to induce green outperformance in the sample including reported and vendor-estimated emissions. However, no statistically significant relationship is found when considering emissions intensity.

Moving to Panel E, we introduce larger lags for emissions data compared to accounting data, based on findings by Zhang (2024) suggesting a median lag of at least 10 months for emissions data. Similar as to Panel E, the results again reveal evidence of green outperformance, but only for vendor-estimated log emissions levels, before scaling by a firm's size. Importantly, the estimate on log emissions levels for vendor-estimated emissions almost equals the estimate in

[1] Given that we have focused on log emissions, a baseline value for log emissions has to be considered in order to interpret the size of the estimated coefficient on log emissions levels. We consider the 20th and 80th percentile to align with the return spread constructed in portfolio analyses, which takes a long position in the value-weighted portfolio based on stocks in the lowest quintile of emissions levels and a short position in the value-weighted portfolio including stocks in the highest quintile in a given month.

[2] These calculations are based on the fact that the 20th quintile of log emissions equals 6.81 while the 80th quintile equals 11.5 for vendor-estimated emissions in the US sample. For the sample incorporating reported and vendor-estimated emissions, these values equal 7.47 and 12.8 respectively.

[3] In particular, the sample period included in the paper by BK2 and Aswani et al. (2023) starts in 2005 already and ends in 2018. Given that our data is from February 2007 onwards, we cannot include data from 2005 and 2006.

Panel E, indicating that accounting for an additional lag in emissions has limited effect.

As previously outlined, the PST framework forms the basis for a negative carbon premium. Specifically, within the PST framework, rapid shifts in investor beliefs could lead preferences towards green firms. In Panel C, we already found that the evidence in favor of green outperformance does not hinge on recent years, which have experienced extreme events, potentially inducing sudden shocks in climate change. We leverage the Unexpected Media Climate Change Concerns (UMC) index developed by Ardia et al. (2022), before the years including these extreme shocks, until June 2018.[4]

Incorporating an interaction term between the UMC factor and emissions measures in the regression, as denoted in Equation 4.5, we find no significant effect on stock returns. Thus, despite a significantly negative coefficient on log emissions suggesting a negative carbon premium, albeit induced by vendor-estimated emissions, we cannot reconcile this effect with the PST model.[5] Although our results diverge from those found by Ardia et al. (2022), who find that brown firms are exposed to heightened climate concerns, we provide a possible explanation for these results. In particular, we find that green outperformance is induced by vendor-estimated emissions. Vendor-estimated emissions may not be considered by investors in their investment decision such that the UMC factor, indicating investor's concern, cannot explain this green outperformance.

The aforementioned findings suggest no evidence of a positive carbon premium and limited evidence of a negative carbon premium, predominantly influenced by vendor-estimated emissions, which diminishes upon adjusting for a firm's size. The largest green outperformance based on log emissions levels is observed in Panel B, before lagging emissions data, implying that a firm with vendor-estimated log emissions levels at the 20th percentile outperforms a similar firm with emissions at the 80th percentile by 2.18% points. Conversely, the lowest green outperformance based on vendor-estimated log emissions levels is found in Panel A, without industry-fixed effects, resulting in a monthly outperformance of 0.83% points. In the sample including reported and vendor-estimated emissions, green outperformance ranges from 1.20% to 1.30% points prior to lagging emissions and accounting data. Recognizing the inherent distinctions between panel regressions and portfolio analyses, we proceed with conducting portfolio analyses.

[4] We recognize that it would be interesting to explore the effect of shocks in climate change concerns in recent years, given extreme events in recent years. However, given that the UMC factor created by Ardia et al. (2022) is available only until June 2018, this is not possible for the current research.

[5] In untabulated analyses, we add the UMC factor to the specifications in Panels D and E of Table 5.1, when lagging emissions data and accounting data. However, the estimated coefficients on the interaction terms between log emissions levels or emissions intensity and the UMC factor are still statistically insignificant. Furthermore, we keep finding evidence of green outperformance for vendor-estimated emissions in Panel D when adding the interaction term and we continue to find no evidence of green outperformance in Panel E, when adding the interaction term.

Table 5.1: Panel regression results based on the US sample (returns in percentages).

	Emissions measure: Log emissions levels			*Emissions measure: Emissions intensity*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Reported & estimated	Reported	Estimated	Reported & estimated	Reported	Estimated
Panel A: US sample, industry-fixed effects not included						
Log emissions levels	-0.096	0.068	-0.177			
	(0.054)*	(0.056)	(0.064)***			
Emissions intensity				0.129	0.237	0.107
				(0.052)**	(0.061)***	(0.061)*
Panel B: industry effects included						
Log emissions levels	-0.243	0.023	-0.465			
	(0.071)***	(0.067)	(0.097)***			
Emissions intensity				0.032	0.075	0.041
				(0.087)	(0.065)	(0.114)
Panel C: US sample, restricted to 2007-2018						
Log emissions levels	-0.227	-0.036	-0.342			
	(0.058)***	(0.060)	(0.084)***			
Emissions intensity				0.033	0.098	-0.046
				(0.076)	(0.050)*	(0.174)
Panel D: US sample, emissions and accounting data lagged by 6 months						
Log emissions levels	-0.115	0.024	-0.237			
	(0.072)	(0.063)	(0.097)**			
Emissions intensity				0.046	0.051	0.055
				(0.088)	(0.069)	(0.109)
Panel E: US sample, emissions and accounting data lagged by 6 and 10 months respectively						
Log emissions levels	-0.120	0.013	-0.234			
	(0.074)	(0.068)	(0.097)**			
Emissions intensity				0.041	0.066	0.045
				(0.094)	(0.070)	(0.113)
Panel F: US sample, including UMC factor						
Log emissions levels	-0.227	-0.005	-0.348			
	(0.066)***	(0.064)	(0.099)***			
Log emissions levels*UMC	0.117	-0.022	0.101			
	(0.131)	(0.120)	(0.154)			
Emissions intensity				0.089	0.113	0.079
				(0.072)	(0.060)*	(0.180)
Emissions intensity*UMC				-0.148	-0.150	-0.335
				(0.268)	(0.247)	(0.427)

Notes: This table presents findings from panel regressions, estimating monthly stock returns in the US market, employing various specifications. The sample period is from February 2007 until January 2023, unless specified otherwise. Panel A showcases the starting point, excluding industry effects, represented as $\text{Ret}_{i,t} = \alpha + \beta \text{Emissions}_{i,t} + \delta \text{Controls}_{i,t} + \mu_t + \epsilon_{i,t}$. Here, $\text{Emissions}_{i,t}$ denotes a generic measure, encompassing either log emissions levels or emissions intensity for firm i at time t. Log emissions levels are measured as $\log(1 + CO_2\text{e emissions})$ and carbon intensity is measured as CO_2e emissions scaled by revenues. All regressions include time-fixed effects, μ_t, and control variables, $\text{Controls}_{i,t}$, although the estimated coefficients are not tabulated in this table. For an overview of included control variables, refer to Table C.6. Panel B extends the analysis by incorporating industry-fixed effects alongside the baseline equation outlined in Panel A. Industry-fixed effects are also included in panels C, D, E and F. Again, the estimated coefficients are not tabulated in this table. In Panel C, the analysis of Panel B is confined to the period 2007-2018 to synchronize with the timeframe of BK. Panel D introduces a lag of six months on emissions data and accounting data. In Panel E, emissions data are lagged by 10 months and accounting data are lagged by 6 months, compared to Panel B. Finally, compared to Panel B, Panel F integrates an interaction term between the UMC factor and emissions, including the UMC factor as formulated by Ardia et al. (2022). Given the limited time frame for which the UMC factor is available, the results in Panel F are based on the sample period from February 2007 until June 2018. Columns (1)-(3) utilize log emissions levels to gauge a firm's environmental performance, while columns (4)-(6) include results based on emissions intensity. Results labeled as 'Reported & estimated' pertain to firms with either reported or vendor-estimated emissions, whereas 'Reported' or 'Estimated' denote findings specific to firms with reported or vendor-estimated emissions, respectively. Standard errors, clustered at the firm and month levels, are reported in parentheses beneath the coefficients in all regressions. Significance levels are indicated by *, **, and *** representing 1%, 5%, and 10% significance, respectively.

5.1.2 Portfolio analysis

This section explores the outcomes of portfolio analyses as outlined in Section 4.2. Figure 5.1 illustrates the cumulative return of the monthly constructed GMB (Green-Minus-Brown) portfolios. By December 2022[6], the cumulative returns of these portfolios amount to approximately 218%, 106%, and 207% for samples comprising reported and vendor-estimated emissions, reported emissions only, and vendor-estimated emissions only, respectively. Examination reveals green portfolios have consistently outperformed brown ones from the middle of 2009 onwards until 2021, particularly evident in the subsample of firms with reported emissions and the combined sample of vendor-estimated and reported emissions. However, for firms with vendor-estimated emissions, a reversal is observed in earlier years, with a cumulative negative spread until mid-2016, after which the cumulative spread remains positive.

Figure 5.2 depicts cumulative returns for the GMB spread based on emissions intensity, revealing differences compared to Figure 5.1. Notably, consistent green outperformance is observed from around 2013 until 2021, particularly in the sample of firms with reported emissions. A similar trend is noted for the combined sample, albeit with outperformance commencing around 2014. However, the cumulative outperformance is relatively lower, approximately 44% for the sample with vendor-estimated and reported emissions. Conversely, no clear green or brown outperformance is discerned when considering vendor-estimated emissions, with negative cumulative returns for much of the period covered.

A remarkable decline in cumulative returns is observed between 2021 and mid-2022, irrespective of sorting by emissions data or emissions intensity. While cumulative returns remain positive for firms sorted by emissions levels, a downturn is observed for those with vendor-estimated emissions when sorted by emissions intensity. This suggests a shift from green to brown outperformance during this turbulent period, marked by events including the Ukraine war and the energy crisis.

Moreover, a notable difference in cumulative return spreads between reported and vendor-estimated emissions is observed. Particularly striking is the similarity in cumulative spread between reported emissions and the combined sample's spread, especially in sorts by emissions levels. This discrepancy may stem from two factors. First, it could be due to the dominance of reported emissions in the combined sample, potentially resulting from a greater volume of observations for reported emissions compared to vendor-estimated emissions. However, examination of Figure B.2 in the Appendix suggests otherwise.

Alternatively, the discrepancy may arise from the value-weighted nature of stocks, potentially favoring firms with reported emissions if these are typically larger market capitalization. In our US sample analysis, firms with reported emissions have an average market capitalization nearly 2.5 times larger than those with vendor-estimated emissions. This bias could affect results when both reported and vendor-estimated emissions are included in the sample, a factor not present in equal-weighted panel regressions.

[6]The original sample includes January 2023, however none of the US companies reporting on emissions in January 2023 have reported both on market capitalization, required for value weighting, and returns.

Figure 5.1: Cumulative returns (%) for Green-Minus-Brown portfolios. Stocks are sorted into quintiles based on emissions levels on a monthly basis, and the return spread results from taking a long position in the value-weighted quintile portfolio with the highest emissions levels and a short position in the value-weighted quintile portfolio with the lowest emissions levels in a given month. We additionally form separate portfolios for stocks with reported emissions and stocks with vendor-estimated emissions. The sample period is from February 2007 until January 2023.

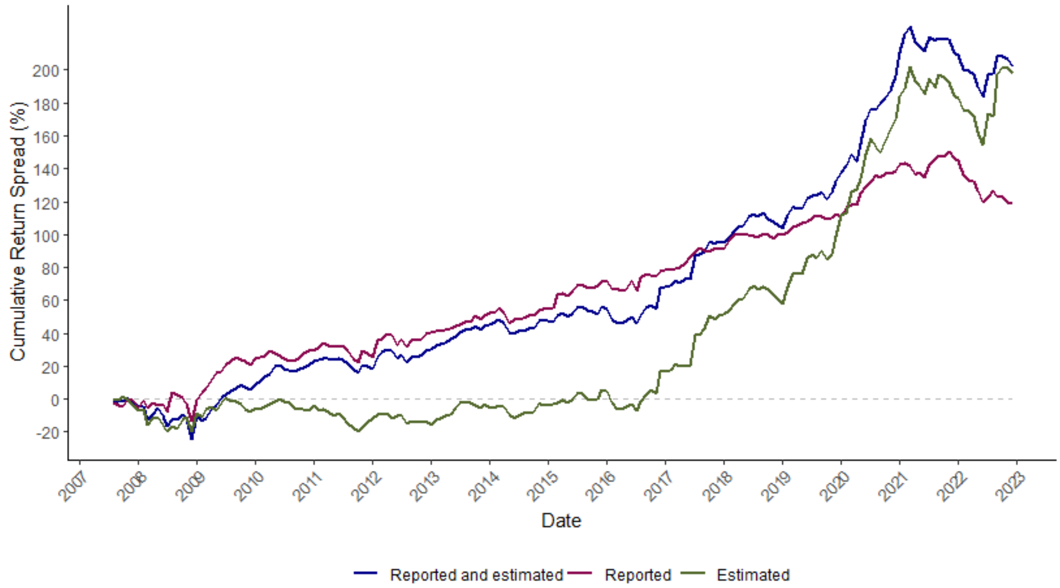
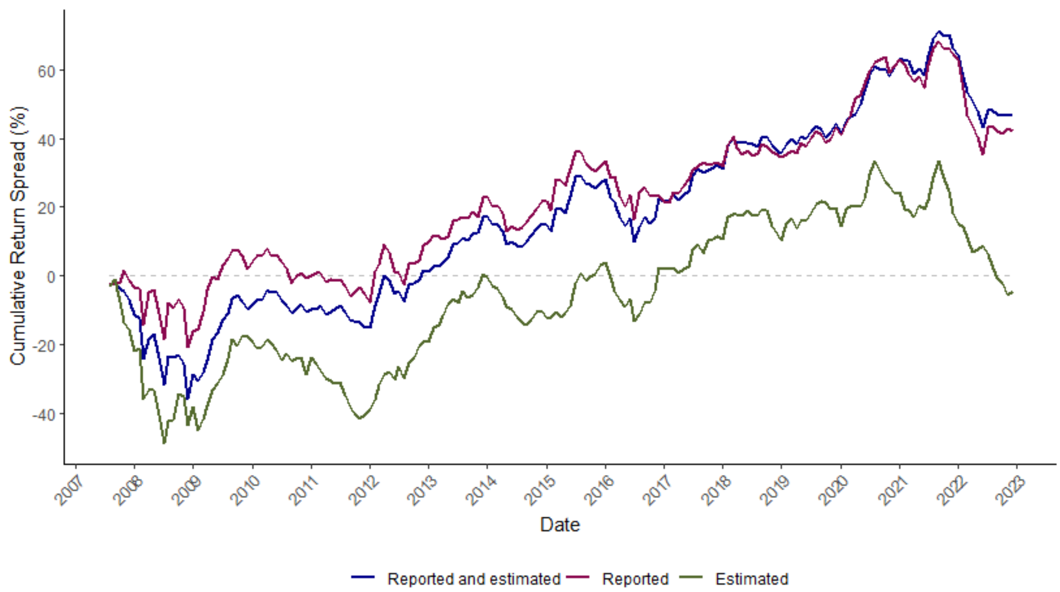


Figure 5.2: Cumulative returns (%) for Green-Minus-Brown portfolios. Stocks are sorted into quintiles based on emissions intensity on a monthly basis, and the return spread results from taking a long position in the value-weighted quintile portfolio with the highest emissions levels and a short position in the value-weighted quintile portfolio with the lowest emissions levels in a given month. We additionally form separate portfolios for stocks with reported emissions and stocks with vendor-estimated emissions. The sample period is from February 2007 until January 2023.



In Panel A of Table 5.2 we report the mean of the value-weighted GMB spread. The GMB portfolio generates a statistically and economically significant positive return of about 1.15% per month, when considering both vendor-estimated and reported emissions and sorting based on emissions levels. These values equal 0.56% and 1.09% for the subsamples including reported or vendor-estimated emissions respectively. Our results are in line with the results of Bauer et al. (2022) who find a return spread equal to 0.45% for reported emissions. However, when sorting based on emissions intensity, we do not find a mean return spread significantly different from zero, for any of the subsamples, in contrast to Bauer et al. (2022) and Aswani et al. (2023), who find significant green outperformance. Similar as to Bauer et al. (2022), we do not find a significant mean GMB spread when performing equal-weighting.

In Panel A, we also document the Sharpe ratios calculated for each of the subsamples when applying value-weighting, based on sorting stocks on either emissions levels or emissions intensity. The monthly Sharpe ratios of the GMB portfolios for the samples in columns (1)-(3) equal 0.234, 0.157 and 0.170. For comparison, the market portfolio's Sharpe ratio equals 0.16 over the sample period.

In Panels B through E, we demonstrate that the observed green outperformance, when sorting based on emissions levels, persists even after controlling for known risk factors and is robust to several modifications. We regress the GMB spread on Fama-French 5 factors and the momentum factor, sorting stocks by emissions levels or emissions intensity, as detailed in Section 4.2. The baseline results, focusing on the intercept capturing a potential premium, are presented in Panel B of Table 5.2. We present the estimated coefficients of the FF5-factors and the momentum factor in Table D.2 of the Online Appendix, for each of the panels in Table 5.2.

Consistent with our prior findings, we identify a statistically and economically significant alpha, the estimated intercept, when sorting based on emissions levels. Notably, this effect is consistent across different samples, including reported and vendor-estimated emissions, reported emissions only, and vendor-estimated emissions only. Again, we find no evidence of green outperformance when sorting based on emissions intensity.

In Panel C, we introduce a 6-month lag on emissions, and in Panel D, we incorporate the UMC factor. Panel E combines both a 6-month lag on emissions and the UMC factor. Considering our findings from portfolio analyses, which indicate green outperformance not solely driven by reported emissions, we may anticipate observing a correlation between the UMC factor and green outperformance. For emissions level-based sorting, we find that green outperformance remains robust to these adjustments, except for vendor-estimated emissions in Panel D, while we find no effect of the UMC factor on the GMB spread. For the combined sample, GMB's alpha remains statistically and economically significant across various specifications, ranging from approximately 1.02% to 1.30%, after accounting for exposure to the FF5 factors and the momentum factor. In untabulated results, we extend the time lag on emissions data to 10 months. However, lagging emissions by 10 months does not alter the conclusions drawn from Panels C and E of Table C.13.

Overall, these findings provide evidence of green outperformance when sorting based on emissions levels, which cannot be explained by unexpected heightened climate change concerns, consistent with our panel regression analysis. Contrary to earlier panel regression findings, portfolio analyses indicate that green outperformance is not solely driven by vendor-estimated emissions. Time lagging has no effect on the conclusions drawn based on portfolio analyses.

Table 5.2: Portfolio analysis results based on the US sample (returns in percentages).

	Sorting based on emissions levels			*Sorting based on emissions intensity*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Estimated & reported	Reported	Estimated	Estimated & reported	Reported	Estimated
Panel A: US sample, without FF5 and momentum						
Value-weighting						
Mean GMB spread	1.146	0.555	1.087	0.229	0.270	−0.183
	(0.406)***	(0.224)**	(0.496)**	(0.250)	(0.221)	(0.275)
Sharpe ratio	0.234	0.157	0.170	0.063	0.071	-0.062
Equal-weighting						
Mean GMB spread	0.199	0.173	−0.048	0.022	0.060	−0.092
	(0.378)	(0.272)	(0.410)	(0.227)	(0.238)	(0.293)
Panel B: US sample, FF5 and momentum						
Alpha	1.300	0.549	1.215	0.231	0.279	−0.173
	(0.381)***	(0.210)***	(0.488)**	(0.221)	(0.219)	(0.306)
Panel C: US sample, FF5 and momentum, 6-month time lag on emissions						
Alpha	1.201	0.653	1.207	0.255	0.226	0.031
	(0.336)***	(0.210)***	(0.420)***	(0.224)	(0.233)	(0.295)
Panel D: US sample, FF5 and momentum, no time lag on emissions, including UMC factor						
Alpha	1.121	0.721	0.709	0.455	0.458	0.430
	(0.303)***	(0.237)***	(0.374)*	(0.258)*	(0.271)*	(0.280)
UMC	−0.204	−0.487	0.143	−0.940	−0.946	−0.616
	(0.872)	(0.825)	(0.935)	(0.843)	(1.010)	(0.841)
Panel E: US sample, FF5 and momentum, 6-month time lag on emissions, including UMC factor						
Alpha	1.016	0.846	0.719	0.471	0.371	0.364
	(0.291)***	(0.235)***	(0.337)**	(0.264)*	(0.232)	(0.316)
UMC	−0.190	−0.277	0.014	−1.178	−0.905	−0.323
	(0.852)	(0.844)	(0.848)	(0.865)	(0.957)	(0.816)

Notes: This table presents the estimated alphas from portfolio analysis on a monthly basis for the US sample. This entails regressing the constructed GMB spread on an intercept, alpha, and in some panels, also on common risk factors. The GMB spread is computed taking a long position in the portfolio consisting of the quintile of greenest firms in a month, measured either by emissions levels or emissions intensity, and a short position in the portfolio consisting of the quintile of the brownest firms in a given month. Emissions levels are measured as CO_2e emissions and emissions intensity is measured as CO_2e emissions scaled by revenues. The sample period is from February 2007 until January 2023, unless specified otherwise. Panel A shows the starting point, represented as $\text{Ret}_t^{\text{GMB}} = \alpha + \epsilon_t$, with t in months. In panel A, the portfolios are either value-weighted or equal-weighted. In Panels B through E, the portfolios are value-weighted. In Panel B, we additionally add common risk factors compared to Panel A, including the FF5-factors and the momentum factor. This results in the following specification: $\text{Ret}_t^{\text{GMB}} = \alpha + \beta_1 \text{RMRF}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{RMW}_t + \beta_5 \text{CMW}_t + \beta_6 \text{UMD}_t + \epsilon_t$. In Panel C, we extend the analysis of Panel B by adding a time lag of 6 months on emissions data. In Panel D, we add the UMC factor as an additional regressor to the regression equation of Panel B. In Panel E, we include both a time lag of 6 months on the emissions measure and add the UMC factor to the regression equation of panel B. In Panels B and C, we only tabulate the estimated coefficient on the intercept, alpha. In Panels D and E, we present both the estimated coefficient on the the intercept and the estimated coefficient on the UMC factor. Columns (1)-(3) utilize emissions levels to sort stocks into quintiles, while columns (4)-(6) include results based on emissions intensity. Results labeled as 'Reported & estimated' pertain to firms with either reported or vendor-estimated emissions, whereas 'Reported' or 'Estimated' denote findings specific to firms with reported or vendor-estimated emissions, respectively. Newey and West (1986, 1994) standard errors are reported in parentheses beneath the coefficients in all regressions. Significance levels are indicated by *, **, and *** representing 1%, 5%, and 10% significance, respectively.

5.1.3 Replicating Bauer et al. (2022)

Thus far, panel regression results suggest a negative carbon premium in the US associated with vendor-estimated emissions when utilizing log emissions levels as an environmental performance metric. However, once emissions are normalized by revenues, no significant premium in either direction emerges. Similarly, portfolio analyses suggest green outperformance when sorting based on emissions levels, yet this advantage is not solely attributable to vendor-estimated emissions.

In Section 4.3, we introduced an indicator variable based on the methodology by Bauer et al. (2022) to bridge the gap between portfolio analyses and panel regression analyses. We highlight several differences between our panel regressions and portfolio analyses. First, portfolio analyses account for systematic risks but not account for firm-specific characteristics and fixed effects, as possible in panel regressions. Moreover, the weighting procedure differs, with portfolio analyses giving more weight to reported emissions than to vendor-estimated emissions due to reported emissions stemming from larger firms on average. Also, portfolio analysis weights each month equally, independent of the number of observations in a given month, while panel regressions give more weight to the months that include more observations by construction. Ultimately, concerns stemming from multicollinearity, which arise from the inclusion of both emissions levels and controls for firm size within panel regressions, do not manifest as issues in portfolio analyses.

Furthermore, we value the method of Bauer et al. (2022), as we noted before, since this is one of the few papers exploring the link between a firm's environmental performance and stock returns while relying on data from LSEG Workspace. In this section, we employ Equation 4.4 as a baseline to explore the link between stock returns and the indicator variable, amending several factors in the specification throughout our analysis.

We tabulate the results based on replicating Bauer et al. (2022) in Table C.11. In each of the panels, we only tabulate the estimated coefficients on the indicator variable. We refer to Table D.3 in the Online Appendix for an overview of the estimated coefficients of the control variables, but not the fixed effects, included in each of the specifications.

In Panel A, we solely regress stock returns on the indicator variable, leaving out the control variables and fixed effects as specified in Equation 4.4. The estimated coefficient on the indicator variable is the closest we can get to calculating the mean returns on the equal-weighted GMB spread, based on panel regressions. Considering Panel A, we find no link between emissions and stock returns, whether relying on emissions levels or intensity. In panel A of Table 5.2 we documented a mean return spread of 0.199% based on emissions levels for the combined sample, an effect also not found to be statistically different form zero. However, in our previous panel regressions we accounted for firm-specific characteristics.

In panel B of Table C.11, we include control variables and industry and time-fixed effects, resulting in the regression given by Equation 4.4. Again, we find that green stocks on average have outperformed brown stocks, by about 0.92% points on a monthly basis. However, this result is induced by vendor-estimated emissions and disappears once we account for a firm's size. In panel C, we amend the equation by adding a time-lag on emissions data of 6 months. The overall conclusion does not change, however the average green outperformance is lowered to about 0.59% points per month. Next, we restrict our sample to the timeframe encompassing 2010-2021 in Panel C, in line with the timeframe employed by Bauer et al. (2022). We continue to find evidence of green outperformance when sorting stocks based on emissions levels but only for vendor-estimated emissions. Our results are partially in line with the results found by Bauer et al. (2022) in their analysis including the indicator variable. In particular, these authors focus on reported emissions and do not find a significant coefficient on the indicator variable when sorting based on emissions levels. However, when including control variables in the regression, the latter authors find evidence of green outperformance while sorting based on emissions intensity, a result that we do not find.

Finally, Bauer et al. (2022) did not account for potential lags in accounting data, only considering lags in emissions data. Given that accounting data have been found to include a lag as well, we present the results when additionally lagging accounting data by 6 months, besides lagging emissions data by 6 months, in Panel E. In Panel F, we account for a potential larger lag in emissions data compared to accounting data, including a log of 10 months on emissions data and 6 months on accounting data. Considering the results, we find no evidence of green outperformance, based on emissions levels or emissions intensity as a measure of firms' greenness, when accounting for potential lags in accounting data. This is not in line with our previous panel regressions, where we continued to find evidence of green outperformance based on log emissions levels.

We note that our overall conclusions are to a large extent in line with our previous conclusions based on panel regressions, with green outperformance induced by vendor-estimated emissions, disappearing once emissions are scaled by revenues. However, relying on the method of Bauer et al. (2022), we find no link between the indicator variable and stock returns, also not for vendor-estimated emissions. Once more, this underscores the sensitivity of outcomes to lagging emissions data and control variables, contingent on the model utilized and whether emissions are assessed continuously or discretely.

5.1.4 Analyses based on uncleaned emissions data

In Section 2.4, we dedicated a significant amount of attention to cleaning the emissions data and in Section 3.2 we already discussed that data cleaning has affected the average emissions levels and standard deviation in our sample. An important question is if the thorough data cleaning process with respect to emissions data has affected our results on the link between a firm's environmental performance and stock returns. To explore this, we replicate panel regressions and portfolio analyses for the US using uncleaned emissions data, pre-cleaning steps detailed in Subsections 2.4.1 to 2.4.8.[7]

Results from panel regressions and portfolio analyses using uncleaned emissions data for the US are tabulated in Table D.6 and Table D.7 in the Online Appendix respectively. Overall, we find similar overall conclusions to Table 5.1. Based on panel regressions, green outperformance induced by vendor-estimated log emissions levels persists, but disappears once emissions are scaled by a firm's size. Time lagging emissions or accounting data continues to result in a carbon premium induced by vendor-estimated emissions, although this effect is not strong enough to result in a negative carbon premium for the sample including reported and vendor-estimated emissions in some of the specifications.

Based on portfolio analyses, we find green outperformance not solely induced by vendor-estimated emissions, which again disappears once we scale emissions by revenues. This effect is robust to lags in emissions data. We find no correlation between heightened climate change concerns and green outperformance, similar as to the cleaned data. However, the vendor-estimated

[7] With the exception that we do lag the estimation method by one month.

coefficients are affected by cleaning the data, both when employing panel regressions and when employing portfolio analyses. For instance, in Panel F of Table 5.1, green firms outperformed brown firms by an average of 1.20% points per month, all else being equal, in contrast to 0.41% points using uncleaned data.[8] This suggests that cleaned and uncleaned emissions data highlight similar sensitivities of conclusions to methodological choices but when interpreting the size of the carbon premium, cleaning emissions data is of importance.

5.2 Global evidence

Our results thus far focus on US firms, in line with much of the prior literature we have discussed. However, as noted in Appendix section A.1, the link between firms' emissions and firms' stock returns may depend on country-specifics. Therefore, in this section we extent our analysis to our global sample, encompassing 90 countries. We follow a similar procedure compared to the US sample, starting with panel regression analyses, followed by portfolio analyses, focusing on a comparison of the results based on the global sample to the results focusing on US firms.

In the panel regressions, where the vendor-estimated coefficients regarding the emissions variables are outlined in Table C.12, the outcomes for the global dataset broadly mirror those observed for the United States. The vendor-estimated coefficients of the remaining control variables for each panel are provided in Table D.4 of the Online Appendix. We discern a green premium associated with vendor-estimated emissions, which diminishes when adjusting for firm size or accounting for data release lags.

Specifically, based on Panel B, we continue to find evidence in favor of a green premium which is induced by vendor-estimated emissions, and disappears once we scale emissions by a firm's size. Similar to the results based on the US sample, lagging log emissions levels and accounting data by 6 months results in green outperformance only for vendor-estimated emissions, with no green outperformance present for the sample encompassing both reported and vendor-estimated emissions. The results are similar when further lagging emissions data to 10 months, although further lagging log emissions levels results in weaker green outperformance. Also for the global sample, we cannot explain the negative carbon premium in terms of shocks to climate change news, as predicted by PST framework. We find no evidence of green outperformance based on emissions intensity. Finally, the size of the coefficient on log emissions levels is found to be smaller for the global sample as compared to the US sample, for each of the subsamples for which we find statistically significant coefficients.[9]

Second, we tabulate the portfolio analysis results for the global sample in Table C.13, comparing these results to the portfolio analysis results for the US as discussed earlier, tabulated in Table 5.2. Again, the estimated coefficients on the control variables for each panel in Table C.13 are presented in the Online Appendix, in Table D.5. Focusing on Panel A of Table

[8]This calculation is based on the fact that the 20th percentile of log emissions levels equals 7.55 for the US sample with uncleaned emissions data while the 80th percentile equals 12.9.

[9]We acknowledge that the difference in size of estimated coefficients between the US and the global sample is not informative in itself when comparing differences in stock returns based on percentiles of emissions given that the percentiles differ between the global and US sample. However, for the US sample we find a difference in log emissions between the 80th percentile of log emissions levels and the 20th equalling 5.33, when considering vendor-estimated and reported emissions. Based on vendor-estimated emissions, this value equals 4.69. For the global sample, these values are 5 and 4.33 respectively. Hence, larger estimated coefficients on log emissions, if significant, in the US sample compared to the global sample, suggest stronger green outperformance in the US relative to the global sample.

C.13, we find a statistically and economically significant average GMB spread of 0.55% for the sample including reported and vendor-estimated emissions, when considering sorting based on vendor-estimated emissions. We note two differences compared to the US results. First, the average spread for the global sample including reported and vendor-estimated emissions is about half the size of the mean GMB spread found for the US. Furthermore, focusing on firms with reported emissions, we find no significant average spread for the global sample, in contradiction to the results based on the US sample. Similar as to the US sample, no green outperformance is found when sorting based on emissions intensity.

In Panels B through E, we account for common risk factors and extend the analysis to account for data release lags in emissions data and we include the UMC factor in panels B through E in Table. In particular, in the global sample, alpha is not found to be significant in any specification considering reported emissions, as opposed to the US sample. Although we continue to find green outperformance for the combined sample in each of the specifications, with estimated alphas ranging from 0.524% points to 0.646% points, this effect cannot be partially attributed to reported emissions. Furthermore, in Panels D and E, when adding the UMC factor, we do find evidence of green outperformance for the combined sample, but this evidence is not evident in samples focusing either or reported on vendor-estimated emissions. That is, the effect found for vendor-estimated emissions is not robust to adding the UMC factor.

Similar as to our conclusions based on panel regressions, we find that the link between stock returns and log emissions levels is stronger in the US sample compared to the global sample, suggesting stronger outperformance in the US compared to the global sample. For the US, the lowest value of alpha results from the specification in panel E of Table 5.2, equaling 1.016% points, while the highest value of alpha for the global sample equals 0.646% in Panel D of Table C.13, being below the lowest value of alpha found for the US sample.

When sorting based on emissions intensity, the estimated intercept is insignificant across panels A through E for each of the subsamples, thus we find no significant (negative) carbon premium that common risk factors cannot explain, similar as to the US results.

Chapter 6

Conclusion

In recent years, understanding the financial implications of firms' environmental performance has gained prominence amidst escalating impacts of climate change and intensified mitigation efforts. The specific focus of our study, the valuation of green firms versus brown firms, has sparked substantial debate within climate finance research, the investment industry, and policy-making. Theoretical perspectives diverge, with classical asset pricing theory predicting a positive carbon premium due to higher regulatory transition risk for brown firms, while the taste argument suggests shifts in consumer and investor preferences could lead to green outperformance. Empirical findings mirror this ambivalence, with studies documenting varied evidence.

In this study, we have undertaken a comprehensive analysis of the link between emissions data and stock returns in an international sample of more than 10,000 listed companies spanning 90 countries, over a period ranging from 2007 to the beginning of 2023. Our analysis sought to explore how methodological choices influence the findings regarding the connection between firms' environmental performance and financial performance. Our analysis aimed to investigate the impact of methodological decisions on the relationship between firms' environmental performance and financial performance. We conducted panel regression and portfolio analyses, compared reported versus vendor-estimated emissions, assessed emissions levels versus emissions intensity, and examined the effects of lagging emissions data and accounting data. These considerations were explored for both the US and global samples.

In the US, our analyses reveal nuanced results. Panel regressions using log emissions show green outperformance, but this effect disappears when focusing solely on reported emissions or scaling emissions by revenues. When lagging emissions levels by 6 or 10 months, and lagging accounting data by 6 months, green outperformance persists for vendor-estimated emissions, but this effect is not strong enough to result in green outperformance in the sample including reported and vendor-estimated emissions. Portfolio analyses also indicate green outperformance in the US, again only for emissions levels, but in this case not solely driven by vendor-estimated emissions. Lagging emissions data by 6 or 10 months does not affect the results. Notably, including the UMC factor does not alter these conclusions and we do not find a link between shocks in climate change concerns and green outperformance, based on either of the models.

Conversely, when examining the global sample, similar patterns emerge but with some distinctions. Panel regressions demonstrate green outperformance, particularly induced by vendor-estimated emissions, consistent with the US findings. Similarly, the effect is not strong enough to result in outperformance for the sample encompassing reported and vendor-estimated emissions, when including lags on emissions and accounting data. Portfolio analyses in the global sample also suggest green outperformance, primarily driven by vendor-estimated emissions levels. However, the inclusion of the UMC factor diminishes the effect for the sample including vendor-estimated emissions only, indicating a more complex relationship on a global scale. Finally, we find no change in results when adding a time lag on emissions data in portfolio analyses. Similar as to the US sample, green outperformance cannot be explained by sudden heightened climate change concerns.

Our study challenges the hypothesis of a positive carbon premium, despite an inability to reconcile the evidence of green outperformance with sudden and unexpected shifts in climate change concerns. Methodological factors significantly influence our conclusions. Specifically, in the US, green outperformance is primarily driven by vendor-estimated log emissions levels in panel regressions, while portfolio analyses demonstrate outperformance even when considering reported log emissions levels alone. In contrast, for the global sample, green outperformance is induced by vendor-estimated log emissions levels in both models. These results underscore the importance of various methodological choices, such as reliance on reported or vendor-estimated emissions, model selection, and the geographic scope of the sample. Moreover, scaling log emissions levels by size provides little evidence of green outperformance, irrespective of these choices. This observation aligns with our finding that emissions variation is largely explained by firm size. This suggests a fundamental link between vendor-estimated emissions levels and stock returns and raises issues due to multicollinearity, as previously suggested by Aswani et al. (2023). The latter issue arises from including emissions levels and the logarithm of market capitalization as regressors within a single regression model. Furthermore, our findings indicate that lagging emissions and accounting data leads to a weakened effect attributed to vendor-estimated emissions, thereby resulting in the absence of green outperformance for the combined sample encompassing reported and vendor-estimated emissions. Regarding portfolio analyses, the temporal lagging of emissions data does not exert an influence on the overarching conclusions.

In conclusion, our findings highlight the continued debate surrounding the link between a firm's environmental performance and its financial performance. Caution is warranted in interpreting this link, as our study reveals nuanced results influenced by various methodological factors. Therefore, further investigation is imperative to identify additional factors contributing to conflicting findings and to elucidate the underlying mechanisms uncovered in this analysis.

Chapter 7

Outlook for future research

Our study highlights the sensitivity of results regarding the relationship between firms' environmental performance and stock returns to methodological choices. Several avenues for future research emerge from our analysis.

First, in this paper we have found conflicting findings on the link between a company's environmental performance and its stock returns, particularly regarding the use of emissions levels or intensity as metrics. While arguments exist for each approach, the preferred measure remains uncertain. Additionally, focusing solely on emissions levels or intensity may overlook firms' strategies for future emissions reduction.Future studies could explore alternative metrics that capture firms' potential for emissions reduction and integrate insights from disclosed emission reduction plans.

Secondly, while we introduced lags for emissions and accounting data, the absence of reported data lags in emissions introduces uncertainty, necessitating detailed information from data vendors to enhance analysis.

In our analyses, through the incorporation of time-fixed effects, industry-specific effects, and country-specific effects into our analysis, we have indirectly accounted for factors related to physical climate risks, which vary both temporally and geographically.[1] Nonetheless, an intriguing avenue for future research could involve investigating the relationship between physical climate risks and stock returns, such as the trend of rising temperatures. In regions such as the Netherlands, risks could even be incorporated at the firm level, given the comprehensive mapping of physical climate risks per region provided by Klimaateffectatlas.[2]

Our baseline results, excluding scaling emissions by firm size or accounting for potential lags, reveal a negative carbon premium induced by vendor-estimated emissions in panel regressions. This contrasts with findings from other papers, including BK's and those by Aswani et al. (2023) and Zhang (2024), prompting questions about the reliance on data vendors. Firstly, coverage varies between vendors, and secondly, different vendors employ distinct estimation techniques for emissions. Aswani et al. (2023) demonstrates that BK's positive carbon premium is induced by vendor-estimated emissions, whereas we find a negative carbon premium induced by vendor-estimated emissions. This suggests that differences in estimation methods across vendors may affect results. In addition, while we distinguish between reported and vendor-estimated data,

[1]However, it is important to note that country-fixed effects may not fully capture physical climate risks at a detailed level, especially in geographically large nations.

[2]For further information on mapped physical climate change risk in the Netherlands, by Klimaateffectatlas, please refer to `https://www.klimaateffectatlas.nl/en/`.

we do not differentiate between the various estimation methods employed by LSEG Workspace. Nonetheless, our analysis in Section 3.5 reveals differences in vendor-estimated emissions between methods, which remain unexplained by firm characteristics. Future research may explore the extent to which results depend on specific estimation methods and the underlying mechanisms involved.

Furthermore, while our conclusions are not affected by cleaning emissions data, we still stress the importance of this process. By manually examining emissions reports and comparing values across periods, we identified significant errors. However, our analysis has limitations, including the absence of comparisons with data from other vendors and insufficient tracking of data updates over time. This underscores the need for stronger regulations governing emissions data collection and reporting, particularly to address firm-level errors. Additionally, empirical research should rigorously clean emissions data to eliminate errors, with cross-validation across vendors recommended, especially for reported emissions. This approach may also reveal discrepancies in vendor-estimated emissions.

In our analysis, we omitted scope 3 emissions due to limited reporting and measurement challenges.[3] However, it is crucial to note that certain sectors, such as mining and gas/oil, may be significantly impacted by this exclusion, potentially biasing our findings. We emphasize the urgent need to improve scope 3 emissions reporting through standardized procedures and stricter regulations. Future research could explore the incorporation of scope 3 emissions into the analysis or investigate their link with stock returns separately, contingent upon improvements in reporting standards.

As a final remark, it is important to acknowledge that our research represents just a fraction of the ever-expanding literature on the implications of firm environmental performance on financial outcomes. With the transition to a net-zero economy gaining significant momentum, we anticipate this research domain will receive even greater attention in the coming decade. While we recognize that our study may not encapsulate every influencing factor on the relationship between a firm's environmental performance and its stock returns, we firmly believe it serves as a crucial cautionary reminder about interpreting statistical associations between carbon emissions and stock returns. We support further research, especially efforts to replicate similar analyses with different data vendors, to better understand this complex relationship.

[3] Another challenge arises from the fact that LSEG Workspace does not offer estimates for scope 3 emissions. Consequently, distinguishing between reported and vendor-estimated emissions would be more challenging when scope 3 emissions are included.

Appendix A

Supplementary information

A.1 Identifying differences across methodologies: in-depth

Section 1.3 sheds light on the conflicting outcomes observed in numerous studies investigating the correlation between a firm's sustainability level and its stock returns. Our study aims to dissect the methodologies employed in this research domain to unravel potential contributors to these disparities. Focusing on previously discussed studies, we delineate six primary areas of deviation, encompassing modeling approaches, metrics for measuring greenness, data sourcing and ESG scores, challenges associated with carbon emissions, temporal inconsistencies, and variations in geographical and temporal scopes. Expanding on the overview provided in Section A.1, this section delves deeper into the nuanced analysis of the divergent methodologies identified across the examined papers.

A.1.1 Modeling choice

First, different studies adopt distinct modeling approaches and Bauer et al. (2022) suggest that methodological choices appear to have a profound influence on the outcomes of a study. For example, BK, who find evidence supporting the carbon premium, employ a regression model adding control variables while Bauer et al. (2022) and Zhang (2024) prefer to rely on portfolio analyses. Important is that these two models are fundamentally different, with panel regressions allowing the option to account for industry-fixed effects, time-fixed effects and firm-specific characteristics, while portfolio analysis cannot account for firm-specific characteristics and instead allows for controlling for common risk factors, such as the Fama-French 5 factors (Fama and French, 1993, 2015).

A.1.2 Metrics on environmental impact of a firm

Second, another notable difference among the studies discussed lies in their chosen metrics for assessing a firm's greenness level. For example, BK, Bauer et al. (2022) and Zhang (2024) concentrate on emissions levels and emissions intensity, while Hsu et al. (2023) aggregate emissions data across various chemical types. Pastor et al. (2022) create a greenness score based on attributes within the 'E' score dataset. It is worth noting that both the choice of greenness scores and the utilization of emissions data have sparked concerns within the literature, as we will discuss later in this section.

A.1.3 Data source and E(SG) scores

Moreover, studies in this field have relied on diverse data vendors, introducing potential challenges for several reasons. One concern arises from the variability in the coverage of firms within different databases. Adding complexity, some studies have utilized ESG scores as a metric for a firm's environmental responsibility. However, this approach has faced criticism due to inconsistencies among data vendors in defining ESG scores or environmental ratings (e.g., Charlin et al., 2022; Berg et al., 2022; La Torre et al., 2020). In addition, investors have voiced more fundamental concerns, questioning whether ESG performance can ever be condensed into a single score (SustainAbility, 2020).

Complicating matters further, Berg et al. (2021) extensively document widespread and recurrent alterations in ESG scores within their sample extracted from LSEG Workspace. They attribute these fluctuations partially to evolving methodologies on which the scores are based and partially to updates in the raw data used for score computation. The authors empirically demonstrate that these data updates influence the outcomes of predictive regressions linking ESG data to future stock returns. Specifically, they observe outperformance in stocks with high ESG ratings when considering updated data, but not when considering the initial data. The authors suggest that this phenomenon stems from data mining, where companies that demonstrated strong performance in a particular year subsequently undergo score upgrades for that year.

Finally, Papadopoulos (2022) comment on discrepancies in evaluating a firm's carbon emissions performance between different data vendors, including MSCI, LSEG Workspace and Urgentem. The authors demonstrate a mixed picture through a simple ranking exercise, which involves ranking firms into five quantiles based on the level of their combined scope 1 and scope emissions. On one hand, in most cases, the rankings align among different data vendors. On the other hand, there are instances where one provider's data would classify a firm as having very low emissions compared to its peers, while another provider's data would categorize it as the most polluting. This suggests that although emissions serve as a more direct measure of a firm's environmental performance compared to environmental scores, relying solely on emissions data when linking a firm's environmental performance to its financial performance may still present challenges.

A.1.4 Issues around emissions data

To tackle the challenges associated with ESG scores, several studies have opted for a narrower focus on subcategories of ESG performance, specifically focusing on emissions levels and emissions intensity (e.g., Bauer et al., 2022; Bolton and Kacperczyk, 2023; Trinks et al., 2022). However, this choice introduces four new potential issues. In particular, we discuss potential biases and inconsistencies in reported emissions data, we discuss the inclusion of different scopes and vendor-estimated emissions in the data, and the impact of using emissions levels or emissions intensity.

First, when examining emissions data sourced from the LSEG Workspace database, Bajic et al. (2023) reveal a series of inaccuracies and inconsistencies.[1] For instance, the authors identify

[1] In this section, we highlight only several examples of inconsistencies found in LSEG Workspace's emissions data. In Section2.4, we provide a much more in-depth analysis of the inaccuracies discussed by Bajic et al. (2023), when performing our own data cleaning.

instances where there is a disparity between the reported total emissions, which are claimed to be the aggregate of scope 1 and 2 emissions, and the actual sum of scope 1 and scope 2 emissions. Furthermore, they find that in some instances LSEG Workspace sets the total of emissions equal to one of the two scopes in case a firm only reported on one of the scopes.

Moreover, as previously noted regarding ESG scores, carbon emissions data is subject to updates over time. In an earlier section, we highlighted the study by Berg et al. (2021), where alterations in ESG scores impacted the identified correlation between ESG scores and stock returns. A portion of these alterations was attributed to amended emissions data by LSEG Workspace. Consequently, the revision of emissions data raises similar concerns. As Bajic et al. (2023) note, the continuous updating of emissions makes it difficult to replicate academic studies using emissions data, once data is updated.

Another point of contention related to emissions data involves the consideration of different scopes. For instance, BK concentrate on carbon emissions levels and their growth, including scope 1, 2, and 3 emissions, while Bauer et al. (2022) and Trinks et al. (2022) do not focus on scope 3 emissions and instead direct attention only to scope 1 and scope 2 emissions because scope 1 and scope 2 are easier to measure, and because disclosure requirements are stricter. Scope 3 emissions are particularly difficult because they lie beyond a company's operational reach and issues such as a wide range of disclosure standards that leave room for interpretation and the potential issue of double counting of emissions across the supply chain, have made reporting difficult (Deloitte, 2023; Shrimali, 2021). As a result, scope 1 data are more consistent across data vendors than scope 2 data, and scope 2 data are more consistent than scope 3 data (Busch et al., 2020).

Fourth, the inclusion of vendor-estimated emissions varies among studies and has been disputed in recent articles. Specifically, data providers can provide estimated emissions based on their own models for companies that did not disclose emissions. For example, BK and Zhang (2024) included firms in their samples with vendor-estimated emissions, thereby assuming the accuracy and compatibility of vendor-estimated data with disclosed emissions.

Aswani et al. (2023) and Busch et al. (2020) criticize the inclusion of vendor-estimated data. First, Busch et al. (2020) find evidence suggesting differences in estimation methodologies across data vendors. Second, Aswani et al. (2023) illustrate that such estimates appear strongly influenced by factors such as firm size, sales growth, industry, and time, rather than capturing variations in environmental impact within industries. Furthermore, they empirically demonstrate that the carbon premium identified by BK1 in US data can be entirely attributed to vendor-estimated emissions data. Aswani et al. (2023) emphasize the importance of their findings regarding the impact of vendor-estimated emissions on the statistical relationship between carbon emissions and returns. While there has been an increase in data coverage, it is primarily driven by vendor-estimated data, making recent data-focused studies susceptible to challenges linked to vendor-estimated emissions.

Recently, the critiques made by Aswani et al. (2023) on the methodological choice by BK to include vendor-estimated data, have been commented on by BK themselves (Bolton and Kacperczyk, 2024). In this commentary, BK offer several reasons for their inclusion of vendor-estimated emissions. First, in an earlier contribution, BK show that despite the growing number of companies disclosing their emissions over the sample period, the average carbon premium also rose. They argue that this simultaneous increase in both the carbon disclosure rate and the carbon premium undermines the claim that the carbon premium is driven by biases in vendor-estimated emissions data (Bolton and Kacperczyk, 2021a). Furthermore, they do find significant evidence of a carbon premium for companies that disclose emissions and attribute this to their sample being about five times larger than the sample of Aswani et al. (2023). Furthermore, they interpret their findings on the carbon premium being lower for firms with disclosed emissions compared to firms with vendor-estimated emissions as evidence of endogenous decisions by firms. Specifically, the smaller carbon premium magnitude observed for disclosed emissions is indicative of a model wherein firms internally determine their decision to disclose emissions. That is, Bolton and Kacperczyk (2024) argue that firms disclosing their emissions require a lower rate of return since these firms are perceived as less risky by investors, compared to firms with vendor-estimated emissions.

In response to the comments made by BK, Aswani et al. (2023) revise parts of their study in a separate comment (Aswani et al., 2024). They note that Trucost, the database that both BK and Aswani et al. (2023) have extracted data from, includes twenty-nine different values for its 'data source' variable and it requires human judgment to determine which values are disclosed or vendor-estimated based on these twenty-nine different values. The main finding in this regard is that BK's result that US disclosed emissions are correlated with stock returns appears to be extremely sensitive to how the disclosure measure is constructed, and fixed effects are chosen. This introduces a concern about the sensitivity and generalizability found by BK concerning the results of disclosed and vendor-estimated emissions.

Moreover, there exists a debate whether to assess emissions in their raw, unscaled form or as emissions intensity, to control for the size of a firm. Advocates of scaled emissions critique unscaled emissions for merging societal reduction goals with individual firms' objectives and suggest that emissions intensity may better align with sustainable investment preferences (Aswani et al., 2023; Nordhaus, 2019). That is, investors who prioritize sustainability may avoid all firms within a relatively polluting industry, such as the oil industry. In line with this argument, Zhang (2024) adds that investors almost exclusively focus on carbon intensity when focusing on net-zero investments. Thus, one can expect emissions intensity instead of emissions levels to be associated with stock returns if investors care about transition risk. Furthermore, Zhang (2024) also shows that emissions are a function of a firm's size and carbon intensity well purges out sales information included in carbon emissions. Finally, Bajic et al. (2023) add an additional argument in favor of using emissions by noting that scaling is commonly based on revenues, and this approach may not be suitable for financially distressed firms or industries marked by volatile production cycles, such as the basic materials sector.

In response to this critique, BK provide several arguments for their choice to focus on unscaled emissions (Bolton and Kacperczyk, 2024). First, they argue that scaling emissions might depict a large company as being more environmentally conscious than a smaller one, despite the larger firm being more polluting. Furthermore, BK argue that intensity is the incorrect measure to use since net zero commitments are based on absolute reduction goals, implying that particularly the largest polluters should prioritize reducing emissions. Finally, they contend that the level of emissions serves as a more direct indicator of carbon transition risk, and accounting for the size of the company by dividing by sales or a related variable introduces noise. That is, according to BK, a change in emissions intensity could then be attributed to a change in emissions levels or a change in the size of the company.

In return, Aswani et al. (2023) write a rejoinder to the comments made by BK and respond to each argument separately (Aswani et al., 2024). We summarize their response. First, the authors argue that it is not an issue that large firms seem more environmentally friendly when focusing on emissions intensity because they argue that forcing firms to reduce consumption is not desirable in case a good is demanded in a specific quantity by consumers. In that case, a greener firm is identified by lower emissions intensity and therefore, emissions intensity would be a more appropriate measure of firm-level carbon risk. Second, (Aswani et al., 2024) note that the argument on net zero pledges holds for emissions levels as well as for emissions intensity. That is, a net-zero pledge can just as easily be interpreted as a pledge to ultimately cut emissions intensity to zero. Furthermore, they argue that such a pledge does not necessarily entail a firm's true intentions to reduce emissions to zero because of the availability of carbon credits to offset existing emissions. Also, (Aswani et al., 2024) acknowledge that scaling emissions by a measure of firm size likely induces noise but recognize this as a less problematic issue than including size as a control variable. In particular, they argue that the latter approach would induce severe multicollinearity because of the mechanical correlation between emissions and returns.

A.1.5 Lagging (emissions) data

Another potential issue regarding emissions data concerns temporal disparities. First, there is often a lag between the conclusion of the emissions reference period and the public release of this data (Bajic et al., 2023; Zhang, 2024, e.g.). This discrepancy is accentuated when aligning emissions data with more promptly reported financial and economic information. In particular, Zhang (2024) argues that the lag on emissions data is larger compared to the lags in accounting data and this is often not properly taken into consideration in papers on the link between emissions and financial performance of a firm. First, proper lags on emissions data are important to align with an investor's knowledge set at a given time. Second, Zhang (2024) argues that the lags adopted for emissions data are often less compared to lags adopted for accounting data. This creates a forward-looking bias for the future accounting data, given that emissions often depend on sales. Zhang (2024) shows that the results found by BK are significantly altered when including proper time lags on emissions data.

Some studies address these temporal disparities, such as Hsu et al. (2023), Bauer et al. (2022) and Zhang (2024), while others, including Aswani et al. (2023) and BK1, study the contemporaneous relation between returns and carbon footprint. BK2 introduce a one-month lag in some of their regressions, not due to the timing inconsistencies mentioned earlier, but because they argue that investors do not immediately react to new information and instead update their information over time.

Second, Bajic et al. (2023) highlight variations in the fiscal year-end dates among different companies while emissions data are usually measured over a calendar year. This can result in the comparison of data from disparate time periods or the aggregation of information spanning multiple calendar years.

A.1.6 Time frame and region

Finally, we observe disparities in the time frames and geographic regions examined across various studies. Recent global events such as the COVID-19 crisis, the Ukraine war, the energy crisis and the ongoing Israeli-Palestinian conflict, prompt inquiry into how these events might impact the research findings associated with these time periods. For instance, an abrupt shift in consumer's and investors expectations during periods of increased climate awareness would anticipate brown stocks' outperformance, as discussed by Pastor et al. (2022) and Pastor et al. (2019). In this light, Meinerding and Zhang (2023) employ textual analysis of newspaper archives and identify shocks in transition risk, including the Paris Agreement in December 2015 and political events such as the bilateral deal between the US and China on carbon emission reductions in November 2014. They find that for these periods, green firms have outperformed brown firms. The difference in time frames is for example clearly visible when comparing the paper by Hsu et al. (2023) to the paper by Bauer et al. (2022). In particular, Hsu et al. (2023) find evidence in favor of the carbon premium in a sample that spans from 1991 until 2016.

On the other hand, Hsu et al. (2023) find evidence of green stocks outperforming brown stocks in their sample from 2010 until 2021. Early years may not be associated with shocks to investor's demand for green and brown stocks. In that case, brown stocks outperforming green stocks in earlier years may dominate any effects from shocks in demand, aligning with the PST framework, potentially explaining the carbon premium found by the former authors. On the other hand, more recent samples may be dominated by the effects from shocks in demand, resulting in a green stocks outperforming brown stocks on average (Bauer et al., 2022).

Furthermore, variations in results may arise when studies focus on different regions. For instance, as demonstrated by Bauer et al. (2022), there are notable differences in the performance of green versus brown stocks across the G7 countries they studied. Most of the studies we have considered so far only focus on the U.S., however, several country-specific factors could potentially influence a firm's environmental performance's impact on its stock returns. For example, it is well established that the level of climate change awareness differs across countries (Bank, 2022; Flynn et al., 2021). In countries where climate change is of less concern to investors, one might expect minimal impact on stock returns. Further, it is reasonable to anticipate that the impact of physical climate risks, which varies among regions and nations, may affect the results. Also, focusing on the US may give an incomplete view given that the financial and regulatory environment, especially in relation to combating climate change, may differ from other countries. Nonetheless, when factoring in country-specific effects, it is crucial to recognize that investors from one nation can invest in stocks from another country, which might alleviate certain country-specific influences such as climate awareness.

A.1.7 In summary

Empirically, the prior discussion underscores the sensitivity of results on the link between a firm's environmental performance and its financial performance to methodological choices. An elaborate overview of methodologies of the studies discussed is provided in Table C.1. Several studies we have reviewed conduct robustness tests or explore variations in their baseline model specifications. While our table aims to offer a comprehensive overview of these methodologies, we do not attempt to cover every single specification. However, we include robustness tests that we deem particularly important, and situations where a study compares two model specifications.

A.2 Elaborate view of data cleaning

A.2.1 Modifying estimation method from yearly to monthly frequency

As discussed, the estimation method used by LSEG Workspace to estimate emissions values in case a firm did not report them, is solely available on an annual basis. Nevertheless, our analysis is on a monthly basis. To align with the monthly frequency of our analysis, we incorporate the estimation method by presuming it to remain constant throughout a firm's fiscal year, covering the 11 months before LSEG reported the firm's estimation method and the month of reporting.

We note however that in some instances the month in which the estimation method was reported changes between a pair of consecutive years for a given year for a firm. As a result, when extending the annual data to monthly data, cases occur where there are two instances of the estimation method for the same month in a year for a given firm.

In case there are two different estimation methods prescribed to a given firm for a given time period due to the extension of the yearly estimation method, we remove the observations for that time period for the firm in question. However, in case the overlap provides two identical estimation methods at a given time point for a certain firm, we keep one of the two observations.

A.2.2 Lagging estimation method: empirical justification

As discussed in Section 2.4.2 we lag the estimation method by one month. For a comprehensive justification of this methodological choice, we compare the estimation method to the variables 'Total CO2 Equivalent Emissions' and 'Estimated Total CO2 Equivalent Emissions', as previously introduced. Since 'Total CO2 Equivalent Emissions' include only reported emissions, in case 'Total CO2 Equivalent Emissions' are filled for a company at a given date, the corresponding estimation method should equal to 'Reported'. There are $519,499$ instances where the variable 'Total CO2 Equivalent Emissions' is filled. Out of these instances, we find $71,098$ instances, including $7,019$ different companies, where the estimation method is unequal to 'Reported', while this number is reduced to $65,158$, including only $3,801$ companies, in case we lag the estimation method by one month. Especially the reduction in the number of companies for which discrepancies between the estimation method and the emissions variables occur is worth to note. That is, if discrepancies occur for multiple periods in a row for a given firm, lagging is not solving the issue. However, for many firms, discrepancies only occur for the first or last fiscal year in a period and lagging the estimation method resolves this issue in many cases. Hence the reduction in the number of companies for which issues occur.

Furthermore, in case the variable 'Estimated Total CO2 Equivalent Emissions' is filled while the variable 'Total CO2 Equivalent Emissions' is not filled, this should imply that the emission value was estimated by LSEG Workspace. That is, the estimation method should equal to 'CO2', 'Median' or 'Energy'. The data contains $471,515$ instances where, based on the two emissions variables, we would expect the estimation method variable to equal to one of these three estimation methods. When we do not lag the estimation method, we identify $59,666$ instances, including $8,283$ different firms, where this is not the case. When we lag the estimation method by one month however, these numbers reduce to $52,896$ and $3,046$ respectively. Based on the above analysis, we lag the estimation method by one month. Thus, there is no estimation method available for January 2007, the initial month of our analysis. Therefore, the sample period is limited to February 2007 through January 2023.

Appendix B

Figures

Figure B.1: This figure provides an overview of the data selection and cleaning process for the global sample, presented with regard to the number of firms and the total number of observations.

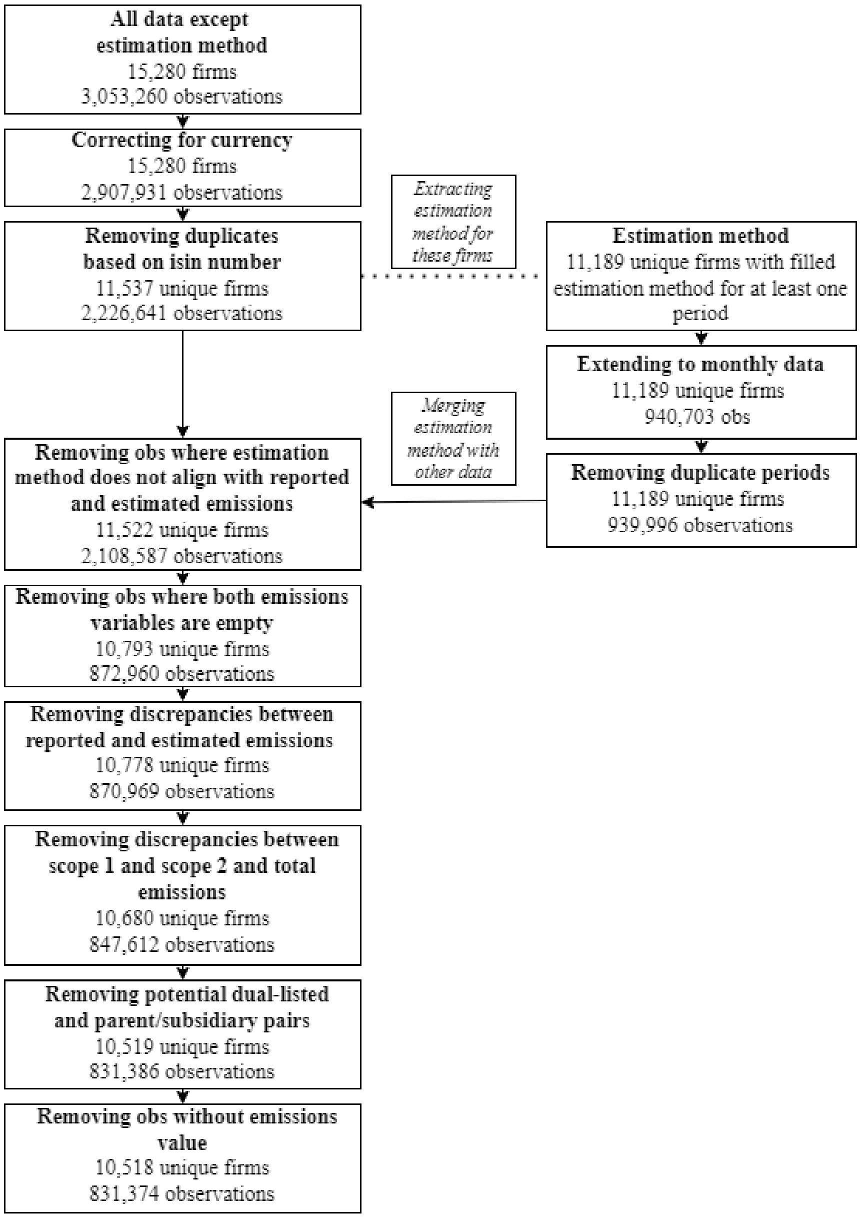
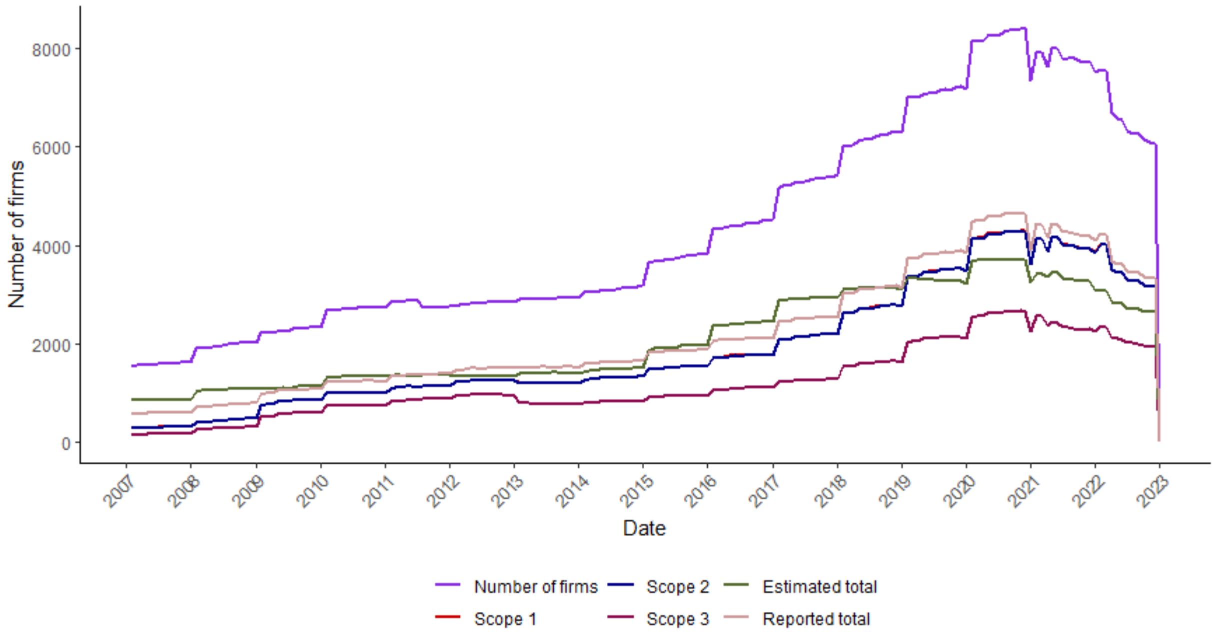


Figure B.2: This figure includes the number of companies with returns data in a given month for the global sample, broken down further into companies with reported or vendor-estimated emissions and data on the separate scopes. The sample period is from February 2007 until January 2023, given that the estimation method is available from February 2007 in our sample.



Appendix C

Tables

Table C.1: This table presents an overview of the (baseline) methodologies employed in different studies investigating the influence of a company's environmental performance on its financial performance.

Paper	**Model**	**Greenness measure**	**Data vendor**	**Emissions data specifics**	**Time frame**	**Geographic Scope**	**Results**	**Additional comments**
Bolton and Kacper-czyk (2021b) (BK1).	Regress environmental measures against monthly stock returns and control variables using pooled OLS.	Emissions levels, emissions growth and emissions intensity (emissions levels scaled by revenues).	Trucost (emissions data), Factset (stock returns).[1]	Separate regressions for scope 1, 2 and 3, and an aggregate of all scopes. Vendor-estimated data are included. Take the log of emissions levels. Do not lag emissions measures in the equation.	2005 - 2017 (yearly).[2]	U.S.	Brown firms outperform green firms, regardless of whether emissions levels or emissions growth are used to assess a firm's environmental sustainability. No significant results are found when focusing on emissions intensity.	Most control variables as well as emissions measures are winsorized.[3] Economic significance of the results increases when including industry-fixed effects. Standard errors are double clustered at firm and year level.
Bolton and Kacper-czyk (2023) (BK2).	Regress environmental measures against monthly stock returns and control variables using pooled OLS.	Emissions levels and emissions growth[4].	Trucost (emissions data), Factset (stock returns).[5]	Separate regressions for scope 1, 2 and 3, and an aggregate of all scopes. Vendor-estimated data are included. Take the log of emissions levels. Emissions measures are lagged by one month.	2005 - 2018 (monthly).	A total of 77 countries were included in the matched samples.	Brown firms outperform green firms, regardless of whether emissions levels or emissions growth are used to assess a firm's environmental sustainability.	Winsorisation is similar as to Bolton and Kacperczyk (2021b). Country-fixed effects, month effects and industry effects are included. Standard errors are double clustered at firm and year level.

Continued on the next page

[1]Control variables are sourced from World Bank, Germanwatch, Morgan Stanley, and IBES.
[2]If the papers include information about the initial month and ending month within the specified time frame, we include this information in our report. However, if the papers do not specify this information, it will not be provided in the table.
[3]For a more thorough explanation of the winsorization process, we recommend consulting the original paper.
[4]The authors only include emissions intensity as a measure of greenness in one of their robustness tests but do not consider it in their baseline models.
[5]Control variables are extracted from databases similar to those utilized by Bolton and Kacperczyk (2021b).

Table C.1: (Continued from previous page)

Paper	Model	Greenness measure	Data vendor	Emissions data specifics	Time frame	Geographic scope	Results	Additional comments
Hsu et al. (2023).	Assess the value-weighted quintile return spread,[6] using their emissions measure to sort stocks into quintile portfolios.	Aggregated emissions for all types of chemicals across all plants listed.	Toxic Release Inventory, Computstat, CRSP.	Emissions are scaled to create various measures of emissions intensity.[7]	1991 - 2016 (yearly).	U.S.	Brown stocks outperform green stocks, across all defined forms of emissions intensity.	Exclude financial firms that have four-digit SIC codes between 6000 and 6999.[8]
Trinks et al. (2022).	Two separate models. Model 1: Assess the value-weighted return spread in a regression including Risk factors,[9] using their emissions measure to sort stocks into decile portfolios.[10] Model 2: Regress environmental measures against monthly CoE and control variables using pooled OLS.[11]	Emissions intensity and sector-adjusted carbon intensity.[12][13]	Refinitiv.[14]	Include the sum of scope 1 and 2 emissions. Exclude vendor-estimated data.	2008 - 2016 (monthly).	A total of 50 countries were included.	Brown stocks outperform green stocks and this effect primarily relates to carbon transition risk (systematic risk) and not to investor's preferences (non-systematic risk).	Country-fixed effects, year effects and industry effects are included in the second model and errors are clustered at the firm level in the second model.

Continued on the next page

[6]The return spread results from taking a long position in the quintile portfolio with the highest level of greenness and a short position in the quintile portfolio with the lowest level of greenness. These portfolios are formed once a year in December, when new emissions data comes in.

[7]Hsu et al. (2023) establish multiple scaling measures, including total assets, property plant and equipment, sales, and market equity, resulting in five separate measures representing the firm's emissions intensity.

[8]These are e.g., finance, insurance, trusts, and real estate sectors.

[9]The risk factors included are global market, size, book-to-market, momentum, profitability, and investment factors (Sharpe, 1966; Fama and French, 1993; Carhart, 2012; Fama and French, 2015).

[10]The return spread results from taking a long position in the decile portfolio of stocks with the lowest level of greenness and a short position in the decile portfolio of stocks with the highest level of greenness.

[11]We refer to the original paper for the formula used to compute the CoE value.

[12]Sector-adjusted carbon intensity is defined as carbon intensity minus the average carbon intensity in the associated sector for a given year, scaled by the standard deviation of carbon intensity.

[13]Emissions are scaled by net sales or revenue.

[14]Robustness analyses are based on emissions data from the Carbon Disclosure Project (CDP) survey. See `https://www.cdp.net/en`.

Table C.1: (Continued from previous page)

Paper	Model	Greenness measure	Data vendor	Emissions data specifics	Time frame	Geographic scope	Results	Additional comments
Bauer et al. (2022).	Primary: Assess the value-weighted monthly quintile return spread,[15] using their emissions measures to sort stocks into quintile portfolios. Secondary: Regress stock returns on an indicator variable indicating if a firm is brown or green and control variables.[16]	Emissions levels and emissions intensity (emissions scaled by total revenues).	Refinitiv.	Only include scope 1 and scope 2 emissions and exclude vendor-estimated data. Account for emissions data to become available in June.	January 2010 - December 2021 (monthly).	G7 countries.	Green stocks outperform brown stocks, particularly considering emissions levels and emissions intensity as measures of greenness.	Apply established criteria to filter out companies with unconventional equity quotations.
Pastor et al. (2019).	Assess the monthly value-weighted tertile return spread,[17] using their greenness measure to sort stocks into portfolios.	Unique environmental score.[18]	MSCI and Media Climate Change Concern Index.[19]	Not applicable.[20]	November 2012 - December 2020 (monthly).	U.S.	Green stocks outperform brown stocks but this outperformance likely reflects an unanticipated increase in environmental concerns.	None.

Continued on the next page

[15]The return spread results from taking a long position in the quintile portfolio with the highest level of greenness and a short position in the quintile portfolio with the lowest level of greenness, with greenness in a certain month determined by emissions data from the year containing this month 18 months back.

[16]The authors define an indicator variable assigned a value plus 1 if a stock falls within the top fifth percentile of emissions in a given month, indicating a brown stock, and minus 1 if it falls within the bottom fifth percentile that same month, indicating a green stock.

[17]The return spread results from taking a long position in the tertile portfolio with the highest level of greenness and a short position in the tertile portfolio with the lowest level of greenness. The score used to define the level of greeness is computed in a certain month based on the most recent data available before this month.

[18]For a comprehensive explanation of the scoring methodology of Bauer et al. (2022), readers are encouraged to consult the authors' original article.

[19]The Media Climate Change Concern Index was constructed by Ardia et al. (2022), using data from eight major U.S. newspapers between January 2003 and June 2018.

[20]This field is relevant only if a study utilizes emissions data, either in the form of emissions levels or emissions intensity.

Table C.1: (Continued from previous page)

Paper	Model	Greenness measure	Data vendor	Emissions data specifics	Time frame	Geographic scope	Results	Additional comments
Zhang (2024).	Baseline: assess the monthly value-weighted quintile return spread,[21] using their emissions measure to sort stocks into portfolios. Secondary: pooled OLS, similar as to BK.	Carbon intensity (emissions levels scaled by fiscal year-end sales).	Same as BK.	Scope 1 and 2 are considered. Vendor-estimated data are included. Emissions are lagged.[22]	June 2009 - December 2021 (monthly).	U.S.	Green stocks outperform brown stocks and the look-ahead bias in the analysis of BK overstates the carbon premium in data.	Control variables are winsorized at 1% and 99% in the regressions. Standard errors are doubly clustered at the firm and monthly level. Carbon measures are standardized to have zero mean and unit variance. Fixed time effects are included.
Aswani et al. (2023).	Baseline: replicated version of BK1.	Emissions levels and emissions intensity (emissions levels scaled by a firm's revenue).	Similar to BK1.	Consider scope 1, 2 and 3 emissions. Make a distinction between estimated and disclosed emissions.	2005-2019.	U.S.	Brown stocks outperform green stocks in case of using emissions levels, based on a similar methodology as BK. However, these results weaken or are not present when the authors consider variations between vendor-estimated and firm-disclosed emissions.	Control variables are winsorized at different values.[23] Fixed time effects are included. Standard errors are doubly clustered at the firm and monthly level. Including fixed industry effects affects the results.

[21]The return spread results from taking a long position in the quintile portfolio with the highest level of greenness and a short position in the quintile portfolio with the lowest level of greenness. The formation of the portfolio in a certain month, using the most recently available return data, is matched to stock returns one month later.

[22]For a more ellaborate description of the lagging process, we refer to the original paper.

[23]For a more thorough understanding of the winsorization process, we recommend consulting the paper authored by Aswani et al. (2023).

Table C.2: The number of companies per region included in the analysis, for different steps during the data cleaning process.

Region	**Num of companies, initially**	**Num of companies, after currency correction**	**Num of companies, after removing duplicates**
United States	3657	3592	3592
Europe	2539	2475	2475
Asia Pacific	4134	4134	4134
North America	4067	3992	462
Africa	209	205	205
Latin America	388	383	383
Mena	286	286	286
Sum	15,280	15,067	11,537

Table C.3: Overview of data amendments based on comparing the variables 'Total CO2 Equivalent Emissions' and 'Estimated Total CO2 Equivalent Emissions', classified based on a number of error types.

Firm	**Estimated or reported**	**Error or suspicion thereof**	**Period**	**Source**
Wrong unit of measurement				
Stepan Company	Reported	In FY 2020 and 2021, the reported variable is of order 2 while the estimated variant is of order 5. Based on the reports, the reported variable is a factor 1000 too small. We amend values accordingly.	FY 2020 and 2021	*2022 Sustainability Report: ESG Analyst Download*
GCC, S.A.B. de C.V.	Reported	In FY 2020 and 2021, scope 1 is reported by LSEG Workspace missing a factor of 1000. This is corrected for in the reported emissions variable but not in the estimated emissions. We amend the estimated emissions accordingly.	FY 2021 and 2021	*2022 Sustainability report*
Incorrect updates				
SFL Co. Ltd.	Reported	Values for FY 2020 were updated by the company in FY 2022, due to an update in scope 1. The estimated variable was updated accordingly by LSEG Workspace but the reported variable was not. We update the reported variable accordingly for FY 2020.	FY 2020	*ESG Report 2022*
Error in reporting by LSEG Workspace[24]				
National Grid plc	Reported	For fiscal year 2020 , the reported value, scope 1 and scope 2 do not match the report of 2021. The estimated value does match this value. We amend the reported value, scope 1 and scope 2 accordingly.	FY 2020	*Responsible Business Report 2020/21*
Suspicious cases with lack of information				
Taiwan Cooperative Financial Holding Co., Ltd.	Reported	No information available in reports. All observations are removed.	ALL	na
Grupo Mexico, S.A.B. de C.V.	Reported	This company is a holding company. The corresponding subsidiary, GCC SAB de CV, is also included in the dataset. We remove both companies.[25]	ALL	na

[24]These are cases other than errors in unit of measurement or incorrect updates. We can rectify these cases but cannot identify the root cause behind the error.

[25]We note that we first amended values for GCC SAB de CV in the above. However, after the discovery of these companies being a holding and subsidiary pair, both companies are removed.

Table C.4: Overview of data amendments based on comparing emissions in successive years for each firm in the dataset, classified based on a number of error types.

Firm	Estimated or reported	Error or suspicion thereof	Period	Source
Wrong unit of measurement[26]				
Rio Tinto Ltd.	Reported	Values were reported in megatons by firm and not amended to tons by LSEG Workspace.	FY 2021 - 2022.[27]	*Scope 1, 2 and 3 Emissions Calculation Methodology 2021*
Hamburger Hafen und Logistik AG	Reported	Values were reported in megatons by company and not amended to tons by LSEG Workspace.	FY 2018 - 2020	*Annual report of 2019, Annual report of 2021*
Lawson Inc.	Reported	Values were reported in thousands of tons by company and not amended to tons by LSEG Workspace.	FY 2019	*ESG data on official website*
Edreams Odigeo	Reported	Values were reported in kg by company and not amended to tons by LSEG Workspace.	FY 2019 - 2020	*Integrated annual report of 2021, Integrated annual report of 2023*
GMexico Transportes S.A.B. de C.V.	Reported	Values were reported in megatons by company and not amended to tons by LSEG Workspace.	FY 2022	*Sustainable development report 2022*
Altri SGPS	Reported	Values in FY 2018 were reported in kg by company but are a amended by LSEG Workspace to megatons instead of tons. Values are removed however given emissions only include CO_2 emissions and equivalent gases are not considered.	FY 2018	*Sustainability report 2018*
Sberbank of Russia	Reported	Values were reported in tons by company but LSEG Workspace did not include a factor of thousand.	FY 2019 and 2021.	*Sustainability report of 2022*
Origin Property	Reported	Values were reported in kilo tons by company and not amended to tons by LSEG Workspace	FY 2021	*Sustainability report 2022*
Albemarle Co.	Reported	Values were reported in thousands of tons by company but not amended to tons by LSEG Workspace.	FY 2021 - 2022	*Sustainability report 2021, Sustainability report 2022*
Brandywine Realty Trust	Reported	Values were reported in kilo tons by company but not amended to tons by LSEG Workspace.	FY 2019 - 2022	*CSR Report 2023*
Cabot Co.	Reported	Values were reported in million tons by company but not amended to tons by LSEG Workspace.	FY 2020	*Sustainability Report 2023*

Continued on the next page

[26] In case we identify a wrong unit of measurement, we amend these cases in the data, by changing the variables Total CO2 Equivalent Emissions and Estimated Total CO2 Equivalent Emissions accordingly. If possible, we also amend the scope 1 and scope 2 reported values.

[27] In case an amendment is made to the reported emissions values, we report the years for which an amendment is made.

Table C.4: (Continued from previous page)

Firm	Estimated vs reported	Error in reported value	Period	Source
Suspicious cases with lack of information[28]				
China Foods Ltd.	Estimated and reported[29]	Reported and estimated values range from order 4 to order 7 in some periods. No information available in reports. All observations are removed.	FY 2009 - 2022[30]	na[31]
Brookfield Co.	Estimated and reported	Reported and estimated values range from order 3 to 6. Only values between FY 2018-2022 harmonize with reports. Any observation before FY 2018 are removed.	FY 2007 - 2022	*Annual report of 2020*
Beijing Oriental Yuhong Waterproof Technology	Estimated and reported	Reported values are of order 5 in FY 2022 while estimated values are of order 9 in FY 2022. No information available in reports. All observations are removed.	FY 2019 - 2022	na
China National Nuclear Power Co.	Estimated and reported	Estimated values range from order 7 to order 8 between FY 2017 - 2022, while reported values are of order 4 in FY 2022. No information available in reports. All observations are removed.	FY 2017 - 2022	na
Xiamen Intretech Inc.	Estimated and reported	Estimated values in FY 2019-2020 are of order 4 while reported values are of order 7 in 2021. No information available in reports. All observations are removed.	FY 2019 - 2022	na
Interconnection Electric SA ESP	Estimated and reported	Estimated values in FY 2010-2013 are of order 6 while reported values are of order 3 in FY 2014-2022. Only information available in reports for FY 2019-2020. Values for these years should be a factor 1000 larger than reported, Although the exact values do not match reports. However, we cannot rectify most years. All observations are removed.	FY 2010 - 2020	*Annual Sustainability report 2020*
Sino-Ocean Group Holding	Estimated and reported	Estimated values are of order 7 in FY 2010-2013 while reported values are of order 4 in FY 2014-2016 and of order 5 in 2017-2022. Values do not match with reports. All observations are removed.	FY 2010 - 2020	*Sustainability reports 2015-2022*
PT Jasa Marga (Persero) Tbk	Estimated and reported	Reported values change from order 2 to order 8, from FY 2018 to 2019. Values in FY 2020-2022 are equal to values in report. Reported values in Sustainable report 2019 are not in line with values in FY 2017,2018 and 2019. All observations before FY 2020 are removed.	FY 2012 - 2019	Sustainable report 2019, Sustainable report 2022
Vodafone Idea Ltd.	Estimated and reported	Both reported and estimated values range from order 0 to order 5 in FY 2010-2022. Annual report of 2023 shows that values are reported in millions of tons, however, none of the reported values, despite the difference in measurement unit, align with the reports. All observations are removed.	FY 2010 - 2022	*Annual report of 2023 and 2015*

Continued on the next page

[28]In several cases, emissions values cannot be found back in reports of the company under consideration or the values reported do not match in any way with the values found in reports. Then, we remove all observations on the concerning company or part of the observations, in case some values can be rectified.

[29]We denote the values to be 'Estimated and reported' in case some of the values that are being scrutinized were estimated while other were reported.

[30]In case all observations on a company are removed, we report the entire time period for which information is available on emissions for the company under consideration.

[31]Fiscal Year is shortened as 'FY'. In case no source is found, we indicate this with 'na'.

Table C.4: (Continued from previous page)

Firm	**Estimated vs reported**	**Error in reported value**	**Period**	**Source**
Mediobanca Banca di Credito Financial	Estimated and reported	Estimated values in FY 2007-2010 are of order 3 or 4 while reported in values in FY 2011-2012 are of order 1. Reported values range from order 2 in FY 2014 to order 4 in FY 2022. Reported values in FY 2016-2022 match with reports. Other values cannot be rectified and are removed.	FY 2007 - 2015	*CSR Reports, 2016 - 2022*
Aker ASA	Estimated and reported	Estimated and reported values range from order 3 to order 8. All observations are removed given that this company exists out of multiple sub companies, and this seems to affect emissions in some years. It remains unclear which emissions are included in which years.	FY 2017 - 2022	*Sustainability report 2021, Sustainability report 2022*
Rubnaya Metallurgicheskaya Kompaniya PAO	Estimated and reported	Estimated values in FY 2008-2016 range from order 5 to 7, while reported values in FY 2017 - 2021 range from order 3 to 4. Reported values are not in line with reports and unclear if scope 3 emissions are included. Values in reported are in millions of tons. All observations are removed.	FY 2008 - 2022	*Environment report 2020*
Alibaba Group	Reported and Estimated	All values are of order 5 or 6 except in FY 2019. Reported values are in line with reports, but value for FY 2019 cannot be found. All observations for FY 2019 are removed.	FY 2019	*2021 Alibaba Group ESG Report, 2023 Alibaba Group Carbon Neutrality Action Report*
Intellicheck	Reported and Estimated	Estimated values in FY 2017 are of order 1, reported values in FY 2019 are of order 5. No information available in reports. All observations are removed.	FY 2017 - 2022	na
Semtech Co.	Estimated and reported	Reported values are of order 0 in FY 2016 - 2021 but of order 4 in FY 2022. Reported values in FY 2022 align with report. No information available for other years. Observations before FY 2022 are removed.	FY 2016 - 2021	*Sustainability report 2023*

Table C.5: Illustration of an identified pair of dual-listed companies and two parent/subsidiary pairs.

Company 1	**Company 2**	**Equivalent emissions (Yes/No)**	**Equivalent revenues (Yes/No)**	**Equivalent financial and accounting variables (Yes/No)**
Parent/subsidiary examples				
Heineken Holding NV (NL)	Heineken NV (NL)	Yes	Yes	No
Dual listed companies				
Vermilion Energy Inc. (US)	Verizon Communications Inc.	Yes	Yes	Yes
Rio Tinto Ltd. (AU)	Rio Tinto PLC (GB)	Yes	Yes	No[32]

[32]Rio Tinto is a special case. According to their website, Rio Tinto operates under a dual listed companies (DLC) structure. However, the data does not show equal values for financial and accounting variables for the studied years. After discussion with LSEG Workspace, we conclude that these two companies did report different financial and accounting variables.

Table C.6: Description of variables included in the analyses. Panel regressions include the dependent variable, emissions variables, the control variables and the fixed effects. Portfolio analyses include the dependent variable and the emissions variables, while controlling for systematic risk factors, which have been explained in Section 4.2.

Variable	Winsorization (in %)	Description
Dependent variable		
Returns	-	Monthly stock returns, measured in percentages.
Emissions variables		
Emissions levels	-	An unscaled measure of a firm's environmental performance. Calculated as the total of a firm's scope 1 and scope 2 emissions, on an annual basis. This includes both reported and vendor-estimated values.
Log emissions levels	-	An unscaled measure of a firm's environmental performance. Calculated as Log(1 + emissions levels), on an annual basis. This includes both reported and vendor-estimated values.
Total emissions intensity	-	A scaled measure of a firm's environmental performance. Calculated as the total of scope 1 and scope 2 emissions, scaled by a firm's revenues, on an annual basis. This includes both reported and vendor-estimated values.
Control variables (panel regressions)		
Logsize	2.5[33]	A measure of a firm's size, calculated as the natural logarithm of a firm's market capitalization, on an annual basis. Market capitalization is calculated as price times shares outstanding.
Book to Market	2.5	A firm's book to market ratio, calculated as the book value of a firm divided by its market capitalization, on an annual basis. Book value is calculated as the total amount of assets minus the total amount of liabilities.
Leverage	2.5	A firm's leverage, calculated as a firm's total debt divided by its common equity, on an annual basis.
Return on equity (ROE)	2.5	A firm's earnings performance, calculated as net yearly income divided by the average of last year's and current year's common equity.
Investment to Assets (Invest/A)	2.5	A firm's investment over assets, calculated as a firm's capital expenditure divided by the book value, on an annual basis.
Momentum	2.5	A firm's momentum, calculated as the average of the firm's stock returns over the past 12 months.
Volatility	2.5	A firm's idiosyncratic risk, calculated as the standard deviation of returns over the past 12 months.[34]
Sales growth	2.5	A firm's growth in sales, measured as a firm's current year's net sales or revenues divided by the firm's last year's total net sales or revenues, on an annual basis.
Country	n.a.	A firm's country of domicile, specified by a firm's ISO code.[35]
Industry	n.a.	The industry in which a firm is categorized.[36]

[33] Market capitalization is winsorized, given it is also used for value-weighting stocks.

[34] Given that our primary analysis starts in January 2007, we collect returns from 2006 onwards to be able to calculate a firm's momentum and volatility.

[35] The ISO Code is the international standard for country codes, by the International Organization for Standardization. We utilize the Alpha-2 code, see `https://www.iso.org/iso-3166-country-codes.html`.

[36] The industry of a firm is defined according to the Refinitiv Business Classification (TRBC) scheme. The system offers five levels of hierarchy structure, including 13 economic sectors, 32 business sectors, 61 industry groups, 153 industries and 895 activities. In our analysis, we use level 4: the 153 industries. Details can be requested via: `https://www.lseg.com/en/data-analytics/financial-data/indices/trbc-business-classification`.

Table C.7: Descriptive statistics based on the global sample.

Variable	N	Mean	StDev. (overall)	StDev. (between)	StDev. (within)	Min	Max
Panel A: general statistics							
Emissions and emissions intensity							
Emissions levels	831,314	2,260,712	35,445,468	33,993,225	26,776,468	0.00	7,630,810,000
Log emissions levels	831,314	10.87	3.01	3.02	0.76	0.00	22.76
Emissions intensity	824,844	0.56	21.85	42.01	12.07	0.00	4,093.23
Control variables							
Leverage	824,556	97.87	131.14	112.45	69.43	0.00	615.26
Sales growth	815,289	11.37	25.80	20.55	21.72	−37.08	105.34
Momentum	754,425	1.08	3.35	2.11	3.12	−6.30	9.98
Volatility	771,084	10.88	5.98	5.11	4.22	3.57	29.60
Book to Market	821,451	0.70	0.58	0.53	0.31	0.04	2.63
Invest/A	820,816	0.11	0.16	0.14	0.09	0.0001	0.82
Logsize	821,583	14.81	1.61	1.56	0.48	11.25	18.06
ROE	803,279	7.92	23.62	23.64	14.49	−79.41	58.61
Panel B: Statistics for uncleaned emissions data							
Emissions levels	890,223	2,659,549	50,973,092	42,912,619	39,919,401	0.00	8,368,040,000
Log emissions levels	890,223	10.92	3.04	3.02	0.8	0.00	22.85
Emissions intensity	889,754	0.70	29.94	45.22	19.26	0.00	4,093.23
Panel C: Statistics per estimation method							
Reported emissions							
Emissions levels	428,197	3,421,103	42,675,246	33,130,855	34,179,568	0.00	7,630,810,000
Log emissions levels	428,197	11.89	2.79	2.95	0.49	0.00	22.76
Emissions intensity	424,723	0.55	11.55	16.36	7.83	0.00	1,772.10
Vendor-estimated emissions using the CO2 method							
Emissions levels	25,017	1,426,550	5,559,954	5,732,195	1,023,906	0.00	64,716,700
Log emissions levels	25,017	11.13	2.96	2.89	0.04	0.00	17.99
Emissions intensity	24,838	0.31	1.45	1.5	0.83	0.00	85.26
Vendor-estimated emissions using the Energy method							
Emissions levels	31,822	7,359,857	89,378,626	138,996,065	5,499,776	0.00	4,350,900,000
Log emissions levels	31,822	11.25	3.10	2.86	0.54	0.00	22.19
Emissions intensity	31,586	2.25	68.10	110.49	2.58	0.00	3,573.51
Vendor-estimated emissions using the Median method							
Emissions levels	346,278	417,474	4,512,459	6,115,480	924,147	0.00	394,022,000
Log emissions levels	346,278	9.55	2.75	2.79	0.66	0.00	19.79
Emissions intensity	343,697	0.44	23.54	47.62	5.39	0.00	4,093.23

Notes: This table includes summary statistics for the global sample. Panel A presents general statistics for the full sample, including statistics on emissions data and the control variables employed in regression analyses. Panel B includes statistics for emissions data based on the dataset before the cleaning process conducted on emissions data. Finally, Panel C includes summary statistics on emissions data when separating reported emissions from vendor-estimated emissions, distinguishing between the three estimation method employed by the data vendor. The sample period is from February 2007 until January 2023.

Table C.8: This table presents summary statistics per country in the global sample.

Country	Frequency	Relative frequency (%)	Min year	Log emissions	Intensity
United States	264,819	31.86	2007	10.24	0.24
China	61,539	7.40	2007	11.77	1.57
Japan	52,987	6.37	2007	12.18	0.17
United Kingdom	51,322	6.17	2007	10.33	0.32
Australia	34,274	4.12	2007	9.51	0.85
Canada	29,496	3.55	2007	11.21	2.91
Germany	21,145	2.54	2007	11.52	0.19
Sweden	19,177	2.31	2007	8.94	0.27
India	18,321	2.20	2007	11.63	1.11
France	18,240	2.19	2007	11.63	0.16
Taiwan, Province of China	17,145	2.06	2007	11.90	0.28
Hong Kong	16,689	2.01	2007	12.13	1.51
Switzerland	16,551	1.99	2007	10.12	0.17
South Africa	14,860	1.79	2007	11.73	0.84
Korea, Republic of	13,095	1.58	2007	11.97	0.13
Brazil	12,048	1.45	2007	11.53	0.37
Malaysia	10,844	1.30	2007	10.84	0.53
Italy	9,684	1.16	2007	11.13	0.28
Spain	7,969	0.96	2007	11.45	0.27
Thailand	7,521	0.90	2007	11.55	0.74
Singapore	7,284	0.88	2007	11.25	0.24
Netherlands	7,126	0.86	2007	11.36	0.11
Turkey	6,900	0.83	2008	11.82	1.27
Mexico	6,348	0.76	2007	11.90	0.26
Denmark	6,225	0.75	2007	10.45	0.12
Finland	6,132	0.74	2007	10.99	0.18
Ireland	6,080	0.73	2007	11.56	0.17
Norway	5,943	0.71	2007	10.17	0.21
Belgium	5,373	0.65	2007	10.64	0.37
Indonesia	5,184	0.62	2008	11.73	0.99
New Zealand	4,891	0.59	2007	9.67	0.37
Russian Federation	4,596	0.55	2007	14.41	1.65
Chile	4,524	0.54	2007	12.22	0.68
Poland	4,217	0.51	2008	11.64	0.74
Bermuda	3,820	0.46	2007	9.53	0.34
Austria	3,615	0.43	2007	12.12	0.23
Argentina	3,438	0.41	2015	10.58	0.73
Greece	3,312	0.40	2007	11.06	0.96
Saudi Arabia	3,312	0.40	2007	11.82	1.31
Philippines	3,216	0.39	2009	12.28	0.56
Israel	3,192	0.38	2007	10.60	0.12
United Arab Emirates	3,012	0.36	2008	9.58	0.72
Peru	2,400	0.29	2008	10.85	0.83
Luxembourg	2,376	0.29	2007	11.64	0.45
Qatar	2,352	0.28	2008	9.62	0.24
Colombia	1,956	0.24	2008	11.38	0.46
Egypt	1,644	0.20	2008	10.24	0.18
Portugal	1,632	0.20	2007	12.15	0.92
Morocco	1,368	0.16	2007	9.00	0.11
Kuwait	1,128	0.14	2008	8.74	0.06
Guernsey	1,080	0.13	2011	7.05	0.68
Cayman Islands	948	0.11	2008	9.89	0.24
Oman	912	0.11	2009	9.86	0.94
Bahrain	828	0.10	2015	7.36	0.11
Hungary	720	0.09	2008	12.24	0.14

Continued on the next page

Table C.8: (Continued from previous page)

Country	**Frequency**	**Relative frequency (%)**	**Min year**	**Log emissions**	**Intensity**
Cyprus	528	0.06	2008	10.06	0.08
Viet Nam	528	0.06	2019	10.97	0.11
Jersey	504	0.06	2007	9.88	0.21
Czech Republic	420	0.05	2007	12.85	1.39
Malta	348	0.04	2015	7.73	0.03
Pakistan	348	0.04	2016	10.81	1.14
Puerto Rico	336	0.04	2008	9.02	0.01
Jordan	312	0.04	2008	8.68	0.03
Macao	300	0.04	2010	12.53	0.11
Iceland	288	0.03	2020	8.68	0.07
Panama	252	0.03	2009	11.71	0.78
Isle of Man	228	0.03	2014	9.47	0.03
Monaco	216	0.03	2016	13.13	2.52
Kazakhstan	192	0.02	2017	12.90	0.62
Romania	192	0.02	2017	11.31	0.91
Sri Lanka	156	0.02	2009	12.17	0.97
Ukraine	156	0.02	2009	13.45	0.33
Uruguay	144	0.02	2016	10.98	0.04
Mauritius	132	0.02	2011	4.98	0.17
Nigeria	120	0.01	2020	11.11	1.00
Slovenia	120	0.01	2017	10.40	0.03
Faroe Islands	96	0.01	2018	9.38	0.09
Kenya	84	0.01	2014	11.00	0.03
Liechtenstein	72	0.01	2018	6.68	0.002
Bahamas	60	0.01	2018	7.13	0.02
Uganda	60	0.01	2018	8.66	0.02
Azerbaijan	48	0.01	2019	10.77	0.52
Cambodia	48	0.01	2019	9.90	0.04
Slovakia	48	0.01	2021	8.27	0.01
Virgin Islands, British	48	0.01	2019	10.94	0.72
Barbados	24	0.003	2019	8.31	0.002
Bulgaria	24	0.003	2021	4.77	0.01
Costa Rica	24	0.003	2019	6.77	0.03
Lebanon	24	0.003	2020	9.34	0.01
Mongolia	24	0.003	2021	12.91	1.25

Notes: This table includes the number of observations per country, the relative frequency of observations per country with the respect to the full sample, the minimal year for which observations exist for a country, and the average of log emissions and emissions intensity, per country.

Table C.9: Panel regression results when regressing measures of firms' emissions on the estimation method and accounting variables, based on the US and global sample.

	Dependent variable:			
	Log emissions levels		Emissions intensity	
	(1)	(2)	(3)	(4)
	US	Global	US	Global
Estimation method: 'CO2'	-0.240	-0.403	0.028	-0.214
	(0.226)	(0.103)***	(0.094)	(0.131)
Estimation method: 'Energy'	-0.203	-0.008	-0.043	1.445
	(0.135)	(0.066)	(0.064)	(1.499)
Estimation method: 'Median'	-0.085	-0.059	-0.048	-0.355
	(0.055)	(0.032)*	(0.023)**	(0.223)
log(Revenue)	0.787	0.875		
	(0.022)***	(0.018)***		
Leverage	0.00003	-0.001	-0.0001	-0.0001
	(0.0002)	(0.0001)***	(0.0002)	(0.001)
ROE	-0.002	-0.003	-0.002	-0.005
	(0.001)***	(0.001)***	(0.001)*	(0.004)
Sales growth	-0.003	-0.002	-0.001	0.006
	(0.0004)***	(0.0005)***	(0.0003)**	(0.009)
Momentum	-0.004	-0.002	0.002	0.011
	(0.003)	(0.003)	(0.001)**	(0.021)
Volatility	0.001	-0.00002	0.002	-0.009
	(0.003)	(0.002)	(0.002)	(0.011)
Logsize	0.194	0.139	-0.016	-0.141
	(0.023)***	(0.019)***	(0.009)*	(0.075)*
Book to Market	0.288	0.307	-0.052	0.286
	(0.049)***	(0.032)***	(0.065)	(0.308)
Invest/A	0.896	1.715	0.140	0.407
	(0.133)***	(0.113)***	(0.079)*	(1.228)
Constant	1.197	-0.575	3.270	4.840
	(0.352)***	(0.269)**	(0.592)***	(2.054)**
Industry-fixed effects	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes
Country-fixed effects	No	Yes	No	Yes
Observations	207,506	722,384	207,506	722,384
R^2	0.906	0.816	0.377	0.031

Notes: This table presents findings from panel regressions in the US and global market, regressing measures of firms' emissions on the estimation method employed for emissions, and accounting variables, on a monthly basis. The sample period is from February 2007 until January 2023. The regression equation is given by $\text{Emissions}_{i,t} = \beta_0 + \beta_1 \text{CO2}_{i,t} + \beta_2 \text{Energy}_{i,t} + \beta_3 \text{Median}_{i,t} + \delta \text{Controls}_{i,t} + \lambda_{industry} + \gamma_t + \epsilon_{i,t}$. Here, $\text{Emissions}_{i,t}$ denotes a generic measure, encompassing either log emissions levels or emissions intensity. Log emissions levels are measured as $\log(1 + CO_2\text{e emissions})$ and carbon intensity is measured as CO_2e emissions scaled by revenues. $\text{CO2}_{i,t}$, $\text{Energy}_{i,t}$, and $\text{Median}_{i,t}$ are indicators denoting whether the estimation method for a given firm i at time t is 'CO2', 'Energy', or 'Median', respectively. The control group comprises emissions that are reported, and firm characteristics for firm i at time t are controlled for through $\text{Controls}_{i,t}$. For additional information on the control variables, see Table C.6. Additionally, industry fixed effects, $\lambda_{\text{industry}}$, and month-year fixed effects, γ_t, are included. In columns (2) and (4), based on the global sample, the regression additionally includes country-fixed effects, country_i. In Columns (1) and (2), the dependent variable is log emissions levels while the dependent variable is emissions intensity in columns (3) and (4). Standard errors, clustered at the firm and month levels, are reported in parentheses beneath the coefficients in all regressions. Significance levels are indicated by *, **, and *** representing 1%, 5%, and 10% significance, respectively.

Table C.10: Panel regression results when regressing log emissions levels on log revenues, for the US sample and the global sample.

	Dependent variable: Log emissions levels					
	(1)	(2)	(3)	(4)	(5)	(6)
	Reported & estimated	Reported	Estimated	Reported & estimated	Reported	Estimated
Panel A: US sample						
log(Revenue)	1.053^{***}	1.008^{***}	0.960^{***}	0.929^{***}	1.000^{***}	0.874^{***}
	(0.015)	(0.046)	(0.015)	(0.011)	(0.037)	(0.013)
Constant	-4.325^{***}	-3.175^{***}	-3.315^{***}	1.454^{***}	0.415	2.273^{***}
	(0.185)	(0.713)	(0.180)	(0.481)	(0.990)	(0.376)
Industry-fixed effects	No	No	No	Yes	Yes	Yes
Time-fixed effects	No	No	No	Yes	Yes	Yes
Country-fixed effects	No	No	No	No	No	No
Observations	262,728	80,071	182,657	262,728	80,071	182,657
R^2	0.644	0.325	0.631	0.902	0.792	0.911
Panel B: global sample						
log(Revenue)	1.034^{***}	1.000^{***}	0.991^{***}	0.965^{***}	0.993^{***}	0.908^{***}
	(0.011)	(0.021)	(0.012)	(0.009)	(0.016)	(0.011)
Constant	-3.810^{***}	-3.132^{***}	-3.408^{***}	0.390^{*}	-0.330	1.355^{***}
	(0.156)	(0.316)	(0.160)	(0.229)	(0.441)	(0.272)
Industry-fixed effects	No	No	No	Yes	Yes	Yes
Time-fixed effects	No	No	No	Yes	Yes	Yes
Country-fixed effects	No	No	No	Yes	Yes	Yes
Observations	824,844	424,723	400,121	824,844	424,723	400,121
R^2	0.523	0.359	0.565	0.816	0.738	0.864

Notes: This table presents the results from panel regressions, estimating monthly log emissions levels by monthly log revenues or log market capitalization in the global market. The sample period is from February 2007 until January 2023. The estimated equation in Panel A is $\text{Log(revenues)}_{i,t} = \beta \log(1 + \text{CO}_2\text{e emissions})_{i,t} + \epsilon_{i,t}$. $\text{Log(revenues)}_{i,t}$ is the log revenue of firm i at time t and $\log(1 + \text{CO}_2\text{e emissions})_{i,t}$ is the log of the emissions levels of firm i at time t. Columns (1)-(3) do not account for industry-fixed effects or time-fixed effects in the estimation while these are included in columns (4)-(6). Furthermore, Panel A is based on the US sample, while Panel B includes results based on the global sample. For Panel B, columns (4)-(6) additionally include country-fixed effects. Results labeled as 'Reported & estimated' pertain to firms with either reported or vendor-estimated emissions, whereas 'Reported' or 'Estimated' denote findings specific to firms with reported or vendor-estimated emissions, respectively. Standard errors, clustered at the firm and month levels, are reported in parentheses beneath the coefficients in all regressions. Significance levels are indicated by *, **, and *** representing 1%, 5%, and 10% significance, respectively.

Table C.11: Panel regression results when including an indicator variable on emissions similar as to Bauer et al. (2022), based on the US sample (returns in percentages).

	Emissions measure: Emissions levels			*Emissions measure: Emissions intensity*		
	(1) Reported & estimated	(2) Reported	(3) Estimated	(4) Reported & estimated	(5) Reported	(6) Estimated emissions
Panel A: US sample, excluding control variables						
Indicator: emissions levels	0.190	0.095	0.175			
	(0.209)	(0.125)	(0.219)			
Indicator: emissions intensity				0.082	0.041	0.011
				(0.112)	(0.138)	(0.146)
Panel B: US sample, including control variables						
Indicator: emissions levels	0.466***	-0.009	0.854***			
	(0.156)	(0.171)	(0.198)			
Indicator: emissions intensity				-0.141	-0.194	-0.262
				(0.129)	(0.150)	(0.178)
Panel C: US sample, including control variables, emissions lagged by 6 months						
Indicator: emissions levels	0.466***	−0.028	0.791***			
	(0.156)	(0.171)	(0.190)			
Indicator: emissions intensity				−0.147	−0.227	−0.040
				(0.136)	(0.150)	(0.178)
Panel D: US sample, including control variables, emissions lagged by 6 months, time frame: 2010-2021						
Indicator: emissions levels	0.543***	0.136	0.868***			
	(0.174)	(0.179)	(0.200)			
Indicator: emissions intensity				−0.021	−0.184	0.127
				(0.135)	(0.165)	(0.164)
Panel E: US sample, including control variables, emissions and accounting data lagged by 6 months						
Indicator: emissions levels	0.170	0.018	0.213			
	(0.164)	(0.163)	(0.215)			
Indicator: emissions intensity				−0.023	−0.122	0.030
				(0.121)	(0.136)	(0.176)
Panel F: US sample, including control variables, emissions and accounting data lagged by 6 and 10 months, resp.						
Indicator: emissions levels	0.170	0.060	0.207			
	(0.164)	(0.171)	(0.211)			
Indicator: emissions intensity				−0.096	−0.125	−0.089
				(0.125)	(0.146)	(0.177)

Notes: This table presents findings from estimating monthly stocks returns in the US market, employing panel regressions and accounting for emissions through an indicator variable. The sample period is from February 2007 until January 2023, unless specified otherwise. In all specifications, an indicator variable indicates if a stock is in the the green portfolio (+1), or in the brown portfolio (−1), or somewhere in between (0), based on quintile portfolio sorts using either emissions levels or intensity, similar as to Bauer et al. (2022). Emissions levels are measured as CO_2e emissions and emissions intensity is measured as CO_2e emissions scaled by revenues. Panel A is based on regressing stock returns on the indicator variable and an intercept. Panel B showcases the findings when accounting for firms-specific characteristics and fixed effects, given by $Ret_{i,t} = \alpha + \beta \text{Emissions indicator}_{i,t} + \delta \text{Controls}_{i,t} + \mu_t + \lambda_{industry} + \epsilon_{i,t}$. $\text{Controls}_{i,t}$ are the firm-specific control variables and time-fixed effects and industry-fixed effects are included via μ_t and $\lambda_{industry}$, respectively. Only the estimated coefficients on the indicator variables are tabulated. Panel C additionally includes the indicator variable based on emissions lagged by 6 months. Compared to panel C, the time frame is restricted to 2010-2021 in panel D. Columns (1)-(3) utilize emissions levels to sort stocks into quintiles, while columns (4)-(6) include results based on emissions intensity. Results labeled as 'Reported & estimated' pertain to firms with either reported or vendor-estimated emissions, whereas 'Reported' or 'Estimated' denote findings specific to firms with reported or vendor-estimated emissions, respectively. In all regressions, standard errors are reported in parentheses beneath the coefficients. In all regressions, *, ** and *** denote significance at the 1%, 5% and 10% levels, respectively.

Table C.12: Panel regression results based on the global sample (returns in percentages).

	Emissions measure: Log emissions levels			*Emissions measure: Emissions intensity*		
	(1) Reported & estimated	(2) Reported	(3) Estimated	(4) Reported & estimated	(5) Reported	(6) Estimated
Panel A: global sample, industry-fixed effects not included						
Log emissions levels	-0.029	0.023	-0.083			
	(0.029)	(0.027)	(0.033)**			
Emissions intensity				0.0005	0.003	0.0002
				(0.001)	(0.003)	(0.001)
Panel B: global sample, industry fixed included						
Log emissions levels	-0.106	-0.024	-0.234			
	(0.029)***	(0.029)	(0.040)***			
Emissions intensity				0.0003	0.002	0.0002
				(0.001)	(0.004)	(0.001)
Panel C: global sample, restricted to 2007-2018						
Log emissions levels	-0.101	-0.044	-0.168			
	(0.024)***	(0.027)*	(0.035)***			
Emissions intensity				-0.004	-0.009	-0.003
				(0.007)	(0.014)	(0.008)
Panel D: global sample, emissions and accounting data lagged by 6 months						
Log emissions levels	-0.040	0.014	-0.128			
	(0.032)	(0.031)	(0.042)***			
Emissions intensity				-0.0002	-0.002	0.0002
				(0.001)	(0.005)	(0.001)
Panel E: global sample, emissions and accounting data lagged by 6 and 10 month respectively						
Log emissions levels	-0.022	0.016	-0.094			
	(0.034)	(0.033)	(0.042)**			
Emissions intensity				-0.0004	-0.001	-0.0001
				(0.001)	(0.004)	(0.001)
Panel F: global sample, including UMC factor						
Log emissions levels	-0.095	-0.031	-0.182			
	(0.027)***	(0.030)	(0.039)***			
Log emissions levels*UMC	0.113	0.043	0.146			
	(0.078)	(0.082)	(0.090)			
Emissions intensity				-0.006	-0.017	-0.002
				(0.008)	(0.019)	(0.010)
Emissions intensity*UMC				0.025	0.024	0.024
				(0.037)	(0.063)	(0.042)

Notes: This table presents findings from panel regressions, estimating monthly stock returns in the global market, employing various specifications. The sample period is from February 2007 until January 2023, unless specified otherwise. Panel A showcases the starting point, excluding industry effects, represented as $\text{Ret}_{i,t} = \alpha + \beta\text{Emissions}_{i,t} + \delta\text{Controls}_{i,t} + \theta\text{country}_i + \mu_t + \epsilon_{i,t}$. Here, $\text{Emissions}_{i,t}$ denotes a generic measure, encompassing either log emissions levels or emissions intensity for firm i at time t. Log emissions levels are measured as $\log(1 + CO_2e \text{ emissions})$ and carbon intensity is measured as CO_2e emissions scaled by revenues. All regressions include time-fixed effects, μ_t, country-fixed effects, $\text{country}_{i,t}$, and control variables, $\text{Controls}_{i,t}$, although the estimated coefficients are not tabulated in this table. For an overview of included control variables, refer to Table C.6. Panel B extends the analysis by incorporating industry-fixed effects alongside the baseline equation outlined in Panel A. Industry-fixed effects are also included in panels C, D, E and F. Again, the estimated coefficients are not tabulated in this table. In Panel C, the analysis of Panel B is confined to the period 2007-2018 to synchronize with the timeframe of BK. Panel D introduces a lag of six months on emissions data and accounting data. In Panel E, emissions data are lagged by 10 months and accounting data are lagged by 6 months, compared to Panel B. Finally, compared to Panel B, Panel F integrates an interaction term between the UMC factor and emissions, including the UMC factor as formulated by Ardia et al. (2022). Given the limited time frame for which the UMC factor is available, the results in Panel F are based on the sample period from February 2007 until June 2018. Columns (1)-(3) utilize log emissions levels to gauge a firm's environmental performance, while columns (4)-(6) include results based on emissions intensity. Results labeled as 'Reported & estimated' pertain to firms with either reported or vendor-estimated emissions, whereas 'Reported' or 'Estimated' denote findings specific to firms with reported or vendor-estimated emissions, respectively. Standard errors, clustered at the firm and month levels, are reported in parentheses beneath the coefficients in all regressions. Significance levels are indicated by *, **, and *** representing 1%, 5%, and 10% significance, respectively.

Table C.13: Portfolio analysis results based on the global sample (returns in percentages).

	Sorting based on emissions levels			*Sorting based on emissions intensity*		
	(1)	(2)	(3)	(4)	(5)	(6)
	Estimated & reported	Reported	Estimated	Estimated & reported	Reported	Estimated
Panel A: global sample, without FF5 and momentum						
Value-weighting						
Mean GMB spread	0.551	0.303	0.702	−0.088	−0.119	−0.138
	(0.242)**	(0.254)	(0.326)**	(0.218)	(0.229)	(0.241)
Sharpe ratio	0.194	0.0933	0.155	-0.0613	-0.065	-0.0683
Equal-weighting						
Mean GMB spread	0.174	0.038	0.103	−0.088	−0.074	−0.053
	(0.185)	(0.148)	(0.230)	(0.184)	(0.169)	(0.228)
Panel B: global sample, FF5 and momentum						
Alpha	0.524	0.302	0.749	−0.079	−0.095	−0.141
	(0.247)**	(0.256)	(0.366)**	(0.229)	(0.272)	(0.229)
Panel C: global sample, FF5 and momentum, 6-month time lag on emissions						
Alpha	0.560	0.240	0.774	−0.063	−0.098	−0.069
	(0.236)**	(0.238)	(0.340)**	(0.212)	(0.237)	(0.236)
Panel D: global sample, FF5 and momentum, no time lag on emissions, including UMC factor						
Alpha	0.646	0.421	0.474	0.226	0.205	0.277
	(0.140)***	(0.228)*	(0.343)	(0.235)	(0.269)	(0.246)
UMC	−0.199	0.009	−0.332	−0.586	−0.704	−0.339
	(0.423)	(0.751)	(0.477)	(0.604)	(0.794)	(0.768)
Panel E: global sample, FF5 and momentum, 6-month time lag on emissions, including UMC factor						
Alpha	0.589	0.311	0.460	0.136	0.147	0.091
	(0.163)***	(0.215)	(0.260)*	(0.249)	(0.271)	(0.295)
UMC	−0.370	0.019	−0.183	−0.602	−0.910	−0.104
	(0.485)	(0.721)	(0.658)	(0.635)	(0.803)	(0.592)

Notes: This table presents the estimated alphas from portfolio analysis on a monthly basis for the global sample. This entails regressing the constructed GMB spread on an intercept, alpha, and in some panels, also on common risk factors. The GMB spread is computed taking a long position in the portfolio consisting of the quintile of greenest firms in a month, measured either by emissions levels or emissions intensity, and a short position in the portfolio consisting of the quintile of the brownest firms in a given month. Emissions levels are measured as CO_2e emissions and emissions intensity is measured as CO_2e emissions scaled by revenues. The sample period is from February 2007 until January 2023, unless specified otherwise. Panel A shows the starting point, represented as $\text{Ret}_t^{\text{GMB}} = \alpha + \epsilon_t$, with t in months. In panel A, the portfolios are either value-weighted or equal-weighted. In Panels B through E, the portfolios are value-weighted. In Panel B, we additionally add common risk factors compared to Panel A, including the FF5-factors and the momentum factor. This results in the following specification: $\text{Ret}_t^{\text{GMB}} = \alpha + \beta_1\text{RMRF}_t + \beta_2\text{SMB}_t + \beta_3\text{HML}_t + \beta_4\text{RMW}_t + \beta_5\text{CMW}_t + \beta_6\text{UMD}_t + \epsilon_t$. In Panel C, we extend the analysis of Panel B by adding a time lag of 6 months on emissions data. In Panel D, we add the UMC factor as an additional regressor to the regression equation of Panel B. In Panel E, we include both a time lag of 6 months on the emissions measure and add the UMC factor to the regression equation of panel B. In Panels B and C, we only tabulate the estimated coefficient on the intercept, alpha. In Panels D and E, we present both the estimated coefficient on the the intercept and the estimated coefficient on the UMC factor. Columns (1)-(3) utilize emissions levels to sort stocks into quintiles, while columns (4)-(6) include results based on emissions intensity. Results labeled as 'Reported & estimated' pertain to firms with either reported or vendor-estimated emissions, whereas 'Reported' or 'Estimated' denote findings specific to firms with reported or vendor-estimated emissions, respectively. Newey and West (1986, 1994) standard errors are reported in parentheses beneath the coefficients in all regressions. Significance levels are indicated by *, **, and *** representing 1%, 5%, and 10% significance, respectively.

Bibliography

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K. (2022). Climate change concerns and the performance of green vs. brown stocks. *Management Science*.

Aswani, J., Raghunandan, A., and Rajgopal, S. (2023). Are carbon emissions associated with stock returns? *Review of Finance*.

Aswani, J., Raghunandan, A., and Rajgopal, S. (2024). Are carbon emissions associated with stock returns?—Reply. *Review of Finance*, 28(1):111–115.

Bajic, A., Kiesel, R., and Hellmich, M. (2023). Handle with care: Challenges in company-level emissions data for assessing financial risks from climate change. *Journal of Climate Finance*, 5.

Bank, E. I. (2022). The EIB climate survey. Report, European Investment Bank.

Bauer, M. D., Huber, D., Rudebusch, G. D., and Wilms, O. (2022). Where is the carbon premium? Global performance of green and brown stocks. *Journal of Climate Finance*, 1.

Berg, F., Fabisik, K., and Sautner, Z. (2021). Is history repeating itself? The (un)predictable past of ESG ratings. *Working paper*.

Berg, F., Kölbel, J. F., and Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26(6):1315–1344.

Bolton, P. and Kacperczyk, M. (2021a). Carbon disclosure and the cost of capital. *Working paper*.

Bolton, P. and Kacperczyk, M. (2021b). Do investors care about carbon risk? *Journal of Financial Economics*, 142(2):517–549.

Bolton, P. and Kacperczyk, M. (2023). Global pricing of carbon-transition risk. *The Journal of Finance*, 78(6):3677–3754.

Bolton, P. and Kacperczyk, M. (2024). Are carbon emissions associated with stock returns? Comment. *Review of Finance*, 28(1):107–109.

Busch, T., Johnson, M., and Pioch, T. (2020). Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology*, 26(1):350–363.

Campiglio, E. and van der Ploeg, F. (2022). Macrofinancial risks of the transition to a low-carbon economy. *Review of Environmental Economics and Policy*, 16(2):173–195.

Carhart, M. M. (2012). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Charlin, V., Cifuentes, A., and Alfaro, J. (2022). ESG ratings: an industry in need of a major overhaul. *Journal of Sustainable Finance & Investment*, pages 1–19.

Deloitte (2023). Challenges in measuring and reporting scope emissions. Report, Ministerie van Infrastructuur en Waterstaat.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Flynn, C., Yamasumi, E., Fisher, S., Snow, D., Grant, Z., and Kirby, M. (2021). Oxford peoples climate vote results. Report, United Nations Development Program, University of Oxford.

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

Giglio, S., Kelly, B., and Stroebel, J. (2021). Climate finance. *Annual Review of Financial Economics*, 13(1):15–36.

Hambel, C. and van der Ploeg, R. (2024). Pricing in transition and physical risks: Carbon premiums and stranded assets. *Working paper.*

Harvard, NPR and RWJF (2022). The impact of extreme weather on views about climate policy in the United States. Report, Harvard School of Public Health, National Public Radio, Robert Wood Johnson Foundation.

Honegger, M. and Reiner, D. (2017). The political economy of negative emissions technologies: consequences for international policy design. *Climate Policy*, 18(3):306–321.

Hsu, P.-H., Li, K., and Tsou, C.-Y. (2023). The pollution premium. *The Journal of Finance*, 78(3):1343–1392.

IPCC (2018). Global warming of 1.5°c. Report, Intergovernmental Panel on Climate Change.

Kabir, M. N., Rahman, S., Rahman, M. A., and Anwar, M. (2021). Carbon emissions and default risk: International evidence from firm-level data. *Economic Modelling*, 103.

Karnani, A. (2011). Doing well by doing good: The grand ilussion. *Californa management review*, 53(2).

La Torre, M., Mango, F., Cafaro, A., and Leo, S. (2020). Does the ESG index affect stock return? Evidence from the Eurostoxx50. *Sustainability*, 12(16).

Loyson, P., Luijendijk, R., and Van Wijnbergen, S. (2023). The pricing of climate transition risk in europe's equity market. *Working paper.*

Matsumura, E. M., Prakash, R., and Vera-Munoz, S. C. (2014). Firm-value effects of carbon emissions and carbon disclosures. *The Accounting Review*, 89(2):695–724.

McGlashan, N. R., Workman, M. H. W., Caldecott, B., and Shah, N. (2012). Negative emissions technologies. Report, Grantham Institute for Climate Change.

Meinerding, Christoph Schüler, Y. S. and Zhang, P. (2023). Shocks to transition risk. Report, Deutsche Bundesbank.

Newey, W. K. and West, K. D. (1986). A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. *Working Paper.*

Newey, W. K. and West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4):631–653.

NGFS (2022). Physical climate risk assessment: Practical lessons for the development of climate scenarios with extreme weather events from emerging markets and developing economies. Report, Network for Greening the Financial System.

Nordhaus, W. (2019). Climate change: The ultimate challenge for economics. *American Economic Review*, 109(6):1991–2014.

Ogunbode, C. A., Doran, R., and Böhm, G. (2019). Exposure to the IPCC special report on 1.5 °C global warming is linked to perceived threat and increased concern about climate change. *Climatic Change*, 158(3-4):361–375.

Papadopoulos, G. (2022). Discrepancies in corporate GHG emissions data and their impact on firm perfo. *Working paper.*

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2019). Sustainable investing in equilibrium. *Working paper.*

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, 146(2).

Sakhel, A. (2017). Corporate climate risk management: Are european companies prepared? *Journal of Cleaner Production*, 165:103–118.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 31(1):119–138.

Shrimali, G. (2021). Scope 3 emissions: measurement and management. *Working paper.*

SustainAbility (2020). Rate the raters 2020: Investor survey and interview results. Report, SustainAbility.

TCFD (2022). 2022 Status Report. Report, Task Force on Climate-Related Financial Disclosure.

Trinks, A., Ibikunle, G., Mulder, M., and ScholtenS, B. (2022). Carbon intensity and the cost of equity capital. *The Energy Journal*, 43(2):181–214.

WBCSD and WRI (2004). A corporate accounting and reporting standard: A revised edition. Report, World Business Councel for Sustainable Developments & World Resources Institute.

Yang, R., Caporin, M., Jiménez-Martin, J.-A., and Kacperczyk, M. (2024). Measuring climate transition risk spillovers. *Review of Finance*, 28(2):447–481.

Zhang, S. (2024). Carbon returns across the globe. *Journal of Finance, forthcoming.*