University of Reading

ESG Finance During Market Downturns

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Finance

ICMA Centre, Henley Business School

Dina Ghanma

May 2022

DECLARATION

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Dina Ghanma

ABSTRACT

This thesis investigates the importance of Environmental, Social, and Governance (ESG) criteria during equity market downturns. Our findings contribute to the debate on whether Corporate Social Responsibility (CSR), as measured by ESG scores, is associated positively with returns. We explore declines that may be accompanied by a shift in investor preference for certain ESG criteria.

The first empirical chapter explores the relationship between European firms' CSR and equity market performance during the Subprime Crisis as well as the Sovereign Debt Crisis (SDC). We find that European firms with higher CSR scores outperformed those with lower scores during the Subprime Crisis. However, we find the opposite association during the SDC and suggest that this may be because CSR scores became more strongly correlated with ESG controversies. We posit that investors have come to rely on more than aggregate CSR measures in assessing a firm's social capital.

Next, we focus on the COVID-19 crisis period and work flexibility – an employee-related measure of ESG under the Social Pillar. We find that firms whose employees were accustomed to work flexibility before the COVID-19 downturn outperformed their counterparts in early 2020. Importantly, we find that this positive association is significantly stronger at the industry level and that it did not exist prior to the pandemic. Our findings contribute to the literature on the potential for socially responsible investments to be a safe haven during market declines.

Finally, we look at market downturn in the countries worst affected by the 2021 summer floods in Europe. Testing a hypothesis that more environmentally responsible firms would outperform their counterparts during this window, we find counter-intuitive results to the contrary. We discuss whether environmental scores may be a proxy for high environmental risk exposure. Our results would fit the supposition that investors suspect ESG scores to be greenwashed.

DEDICATION

I dedicate this thesis to my parents, Riadh and Raja', and my sisters, Samar, Sara, and Tala, for being the best family a person could ask for. My loving parents' support is unrivalled, and my sisters are my best friends. I owe it all to them.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr Ioannis Oikonomou, Professor Simone Varotto, and Dr Lisa Schopohl for their support during my PhD studies and for their helpful comments and insights about my research. I am very grateful to have had such excellent researchers to guide me on my journey. Thank you for the advice and patience along the way.

I am also grateful to the ICMA Centre family for being my 'home away from home' from when I first arrived in the UK to do my BSc in 2013. I am very appreciative to remain a part of this community and would like to thank Professor Carol Padgett and Professor George Alexandridis for their faith in me and the wisdom they have shared. I would also like to thank Professor Adrian Bell, Dr Gita Persand, Professor Mike Clements, Dr Nadia Kappou, and Dr Mike Smith for their leadership when I taught during my studies. I am also grateful to the Centre's administrative staff who are always happy to help. Thanks to Leanne and Karen for adding zest to the building and always making me feel welcome. To all my fellow PhD students across the University of Reading – Ludovico, Fiona, Tommaso, Safa'a, Corina, Victoria, Renato, Robert, Vasilis, and others – thank you for the coffee breaks, the fruitful discussions, and the fun times.

To Liviu, Sara V., Despoina, Zary, Samuele, Massimo, Elisavet, Evripidis, Yazan, Dana, and Omar, thank you for the friendship, support, and words of encouragement throughout the years. I would also like to thank my grandparents and my many aunts, uncles, and cousins who have always been a source of joy and motivation in my life.

My heartfelt appreciation goes to my mentor, Dr Ioannis Oikonomou, for seeing potential in me and getting me started on this path. He was a remarkable educator, brilliant mind, and wonderful person. He was an inspiration to me and many others. I am blessed to have known him. May he rest in peace.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

B/M	Equity Book-to-Market Ratio
CAPM	Capital Asset Pricing Model
CCPI	Climate Change Performance Index
CFTC	Commodity Futures Trading Commission
CFP	Corporate Financial Performance
COP26	2021 United Nations Climate Change Conference
CSR	Corporate Social Responsibility
ESG	Environmental, Social, Governance
EU	European Union
GICS	Global Industry Classification Standard
GIIPS	Greece, Italy, Ireland, Portugal, and Spain
ITUC	International Trade Union Confederation
SDC	Sovereign Debt Crisis
SFDR	Sustainable Finance Disclosure Regulation
SRI	Socially Responsible Investment(-s, -ing)
UK	United Kingdom
UN	United Nations
UNDRR	United Nations Office for Disaster Risk Reduction
UNPRI	United Nations Principles for Responsible Investment
US	United States
WFH	Work(ing) from Home
WHO	World Health Organization

1 INTRODUCTION

For many years, Corporate Social Responsibility (CSR) has been a topic of interest in both academia and industry. How best to define CSR has been one of the popular questions for discussion. A relatively early definition from the European Commission (2001) describes CSR as a "concept whereby companies integrate social and environmental concerns in their business operations and in their interaction with their stakeholders on a voluntary basis".[1] More recently, it has become common practice to use the term CSR interchangeably with sustainable or ethical business and corporate citizenship, according to Business for Social Responsibility™ (Olson, 2013). Indeed, in its ISO26000 guidance on CSR, the International Standards Organisation asserts that "the objective of social responsibility is to contribute to sustainable development" (ISO, 2010).

Given the highly qualitative nature of CSR as a concept, a natural consideration is how best to measure it quantitatively. ESG criteria have increasingly been looked to for this. Lexology (2021) provides a practitioner-contributed characterisation of the relationship between CSR and ESG, stating that "while CSR aims to make a business accountable, ESG criteria make [such efforts] measurable". The same source defines ESG as business indicators used by socially conscious investors to identify and quantify a company's sustainability and societal impact. The close association between ESG, CSR, and social capital has given rise to ESG scores being used to proxy for CSR in numerous empirical studies (e.g., Hong and Kostovetsky, 2012; Krüger, 2015; Lins et al., 2017; Nollet et al., 2016; Yoon et al., 2018 and others).[2]

A lot of work has been done to understand CSR's impact on financial and non-financial aspects of business and community development. While some findings praise CSR as being highly

[1] This has since become popular and remains a common definition adopted by various global bodies and organisations, including the UN Industrial Development Organisation (2022).
[2] For this reason, while recognising that CSR and ESG are distinct, we refer to them interchangeably at times throughout this thesis, as is common in the corporate finance literature (Gillan et al., 2021).

relevant and financially rewarding to those companies which invest in it (e.g., Erhemjamts et al., 2013), others have deemed it irrelevant (e.g., McWilliams and Siegel, 2000) or even harmful (e.g., Peng and Yang, 2014).

Recently, a subset of the literature has focused on the relationship between CSR, often measured using Environmental, Social, and Governance (ESG) indicators, and corporate financial performance (CFP) during market downturns. This could be due to the theorised resilience of socially responsible investments (SRI) to selling pressure. Compared to conventional investors, socially responsible investors are thought to be less sensitive to financial performance (Bollen, 2007; Pástor et al., 2021b; Renneboog et al., 2011). There is evidence that they are slower to penalise poor financial performance because of the utility they derive from holding more ethical investments (Bollen. 2007; Geczy et al., 2003). Alternatively, investors may be financially motivated to consider ESG criteria in their portfolio management; notably, ESG integration has been found to provide meaningful downside risk protection (e.g., Kaiser and Welters, 2019; Sherwood and Pollard, 2017; Verhedyen et al., 2016). As such, when the market is in turmoil, higher ESG-scoring investments are predicted to outperform, especially as investors become more socially responsible.

There can also be more particular reasons to expect the CSR-CFP link to be different during crisis windows. For example, during the Subprime Crisis, Lins et al. (2017) put forth that trust in the corporate sector was low. As such, they hypothesise and show that social capital, proxied by CSR measures, became increasingly connected to better firm equity performance. We build on this in Chapter 3. During the more recent COVID-19 pandemic, it was the social responsiveness of companies to their employees and local communities that was predicted to influence returns (Folger-Laronde et al., 2020). We consider this in Chapter 4. If a market decline is caused by unexpected climate deterioration, ‘green’ assets are expected to outperform ‘brown’ ones (Engle et al., 2020). This is the notion underlying Chapter 5.

The common theme in the empirical chapters of this thesis is the investigation of how ESG measures, either aggregate or single-dimensional, are related to equity market performance during market downturns. Our baseline econometric approach is the use of a cross-sectional regression, where crisis-window buy-and-hold returns are explained by a lagged measure of CSR and controls. We then expand on these findings by exploring how the CSR-CFP link may have changed during the crisis window as compared to previously, using panel regressions. We also make use of interaction variables to investigate the CSR measure's relationship within subsets of firms – for example, in certain sectors.

In the first empirical chapter, we start by methodologically replicating Lins et al.'s (2017) investigation into CSR and US equity market performance for a European sample. Using a cross-sectional regression, we find that high-CSR firms outperformed low-CSR firms in Europe during the Subprime Crisis. However, we find that the opposite is true for the same sample of firms during the Sovereign Debt Crisis (SDC) two years thereafter. To explore the reason behind this, we turned our attention to the CSR scores themselves. We find that CSR scores around the SDC were much more highly correlated with year-ahead ESG controversies. We suggest that this may be a form of 'greenwashing', whereby firms with knowledge of impending ESG-related lawsuits and fines window-dress their CSR measures to appear more responsible than they are. We find that using ESG controversy avoidance as a measure of responsibility in place of aggregate CSR scores reveals the more intuitive positive association between CSR and CFP.

In our second empirical chapter, we focus on a single dimension of CSR – work flexibility. This is a component of the Social Pillar of ESG, under policies capturing a firm's social responsiveness to its workforce. Using a sample of firms domiciled in developed markets, we find that firms which had flexible working policies prior to the COVID-19 pandemic weathered the equity market crash in the first quarter of 2020 better. Our results suggest that the

association between work flexibility and improved performance is strongest at the industry level. We further find that flexible working became significantly more value-adding after the COVID-19 pandemic while it was either negligible or harmful prior to the pandemic, likely because of the perceived direct and indirect costs associated with work flexibility.

Finally, in our third empirical chapter, we turn our attention to the heightening conversation around climate change and environmental responsibility. We consider whether the flash floods in Europe during the summer of 2021 were accompanied by any sentiment-driven investing towards more environmentally responsible companies in four countries materially affected by the floods. We find a counter-intuitive result that shows higher environmental (E) scores being negatively associated with equity market performance during the flooding window. Our findings are not in line with work that predicts green firms outperforming brown ones during unexpected climate deterioration (e.g., Pástor et al., 2021a), noting we use a different methodology, though. When replacing aggregate E-scores with more tangible environmental indicators, like whether a firm makes expenditures on environmental initiatives, we find a positive but insignificant association with flooding-window returns.

The overarching intended contribution of this thesis is to further elaborate on the relationship between CSR and CFP, which remains highly contested. By focusing on three different equity market downturns, we also contribute to the subset of the literature investigating whether SRI can be a safe haven during crisis periods and periods of high volatility. We also contribute to the defining of CSR, another area that has garnered wide academic and practitioner efforts over the years with no clear consensus. This thesis also discusses whether ESG scores may be 'window-dressed' by managers, perhaps through the pursuit of less substantive CSR initiatives (Li et al., 2019). This is by considering the relationship between ESG scores, which are susceptible to managerial influence, and less biased CSR indicators. The latter include media-reported ESG controversies and verifiable information about ESG expenditures. By

considering the form of responsibility expected to be most relevant during a particular market episode, we contribute to the segment of the CSR literature that supports the disentanglement of aggregate ESG measures in favour of the more verifiable indicators. Furthermore, by conducting a large portion of our analyses on European markets, we expand the still US-centric research on CSR (Pisani et al., 2017) and particularly 'green' investments (e.g., Pástor et al., 2021a) to wider geographies.

To summarise, our work in this thesis emphasises the limitations of ESG scores and the continued lack of consensus on the relationship between CSR and financial performance. Areas for future work include the disentangling of aggregate ESG scores to more verifiable CSR dimensions, as well as gaining a better understanding of which CSR dimensions are most contemporaneously relevant at a given time. As investors become more concerned with matters like sustainability, equality, and diversity, it becomes increasingly vital that corporate-level measures of social responsibility become more reliable and consistent.

The remainder of this thesis is structured as follows. Chapter 2 gives an overview of the relevant literature, which is expanded upon in each of the empirical chapters – Chapters 3, 4, and 5. For each of these, we start by introducing the chapter and summarising the relevant literature for the hypotheses tested and methodologies used. We then summarise our sample statistics, present our econometric approach, and discuss our results. We finally conclude each chapter by summarising our findings and discussing areas for future work. Chapter 6 offers the conclusions for the thesis overall. The references cited throughout the thesis are included at the end.

2 LITERATURE REVIEW

Many classical economists, like Friedman (1970), were critics of the notion that corporations have a responsibility to stakeholders beyond their investors. Their view was opposed by Davis (1973) and others who believed that companies ought to consider the consequences of their actions since economic and legal requirements were recklessly narrow. The earliest work on CSR was largely an ethical discourse, but soon after, the more popular question became whether companies can 'do well by doing good', often referred to as the 'business case for CSR'. Empirical work on the relationship between CSR and CFP is plentiful but reaches no consensus. Findings of a positive (e.g., Erthemjamts et al., 2013; Rodgers et al., 2013), negative (e.g., Bolton and Kacperczyk, 2021; Peng and Yang, 2014), non-existent (e.g., Hoepner and Schopohl, 2018; Lindsey et al., 2021; McWilliams and Siegel, 2000; Sun et al., 2010), and curvilinear (e.g., Barnett and Saloman, 2006, 2012) relationship have all been put forth.

To clarify the CSR-CFP relationship, the reasons behind the findings on its nature being so sporadic must first be understood. Then, if possible, the limitations clouding the true CSR-CFP link must be addressed. Numerous studies have offered explanations for why this relationship is so difficult to contain (Galent and Cadez, 2014; Surroca et al., 2010). These include concerns over endogeneity, how best to measure CSR, and the legitimacy of CSR reporting. We briefly discuss each of these in this review.

First, it is notoriously difficult to determine the direction of causality between CSR and firm profitability (e.g., Waddok and Graves, 1997). During normal times, it is difficult to say whether the line of causality goes from a firm's regard for stakeholder welfare to its market performance or the opposite. It is very plausible that the firms 'doing well' are the ones with the capacity to 'do good'. However, recent work has mitigated this by observing 'natural experiments'. For example, Lins et al. (2017) circumvent the issue by focusing on a period when an exogenous financial shock disrupts the equilibrium in the absence of short-term

changes to CSR ratings. The use of such a setting has also been adopted for more recent crises, like the 2020 COVID-19 crash (e.g., Bae et al., 2021). In the same way, the crises we investigate in this thesis serve as natural experiments for testing the CSR-CFP relationship, allowing us to make inferences about the value of CSR during an unexpected market downturn.

Another key issue in CSR research is how best to define and measure CSR. There is no strong theoretical foundation for CSR as a concept (Crane et al., 2008; Dahlsrud, 2008; Sheehy, 2015). Reputation indices remain the most popular way of measuring CSR (Diez-Cañamero et al., 2020; Galant and Cadez, 2014). These include the ratings compiled by MSCI (previously KLD), Refinitiv (sometimes referred to as Thomson Reuters, and previously ASSET4), and Sustainalytics. MSCI ratings may be the most comprehensive among these and are likely the ones primarily used in CSR research (Coombs and Gilley, 2005). However, for many years, this database only rated companies in the US. This made it difficult to test the same CSR relationships in other regions without introducing inconsistencies from an alternative CSR measurement. Even now, for historical studies, the results may be affected by the retrospective assignment of these CSR ratings for non-US firms. While other reputation indices existed for regions beyond the US, they often had limited firm coverage. This is one of the reasons the CSR-CFP relationship was often studied in US companies only, limiting true understanding of this relationship (Adam and Shavit, 2008). However, Refinitiv's ESG database had been gaining popularity for its international scope and is now widely used for CSR research (e.g., Auer and Schumacher, 2016; Bae et al., 2021; Cheng et al., 2014; Fernandez-Feijoo et al., 2014; Jost et al., 2022, and others). In this thesis, we follow these authors and use the Refinitiv ESG database to construct our CSR measures across the three empirical chapters.

Another area of discussion around CSR ratings is the matter of reliability. Until there are clear-cut criteria for what constitutes 'good' versus 'bad' CSR, its measurement will remain inconsistent and subject to bias (Griffin and Mahon, 1997). CSR measurements from these

major providers are said to suffer from sampling limitations which are two-tiered (Van Beurden and Gössling, 2008). Focusing on reputation indices, the first tier is the selection bias from the choice of firm coverage by the ESG rating agencies. Typically, reputation indices, like those of MSCI and Refinitiv, concentrate on large, publicly listed companies. Such firms are more likely to be concerned with CSR and have the means to invest in it. The second tier has to do with the reliance on companies voluntarily disclosing their CSR information, as this was not mandatory in most countries for a long time and remains voluntary in many places. Even where CSR disclosure is mandated, this would not cover all the indicators collected by the ESG ratings providers. As such, intuitively, it is more likely for the firms with better CSR to divulge information relating to their practices, leading to a reporting bias. This would not only imply that reputation indices capture the best tranche of companies, but that the ratings might also be inflated due to companies' impression management (Turker, 2009). However, as CSR disclosures have become more commonplace, more and more companies are being rated by ESG database providers, so some of the concerns around sampling bias have diminished. It is also worth noting that ESG scores from different providers may not necessarily agree on a firm's ESG performance (Berg et al., 2019; Christensen et al., 2021). Berg et al. (2021) find that relying on several complementary ESG ratings instead of ratings from a single provider is advantageous in clarifying the relationship between ESG performance and other variables.

A key to understanding what CSR scores measure is clarifying what form of CSR companies undertake and their reasons for doing so. Naturally, a key motivation for companies to pursue CSR is to attract and retain investors, many of whom are becoming increasingly interested in ESG criteria (Bernow et al., 2017; Boffo and Patalano, 2020; Eccles and Klimenko, 2019). The motivations behind investors seeking responsible investments are also discussed in the literature. Daugaard (2019) provides an overview of the cultural and historical reasons that have given rise to various forms of ESG investing. There is evidence that ethical investors with

major shareholdings may demand that managers behave sustainably and responsibly, regardless of the impact on their gains (e.g., Mackey et al., 2007). Even when early empirical work suggested that ethical investments do not outperform conventional ones (e.g., Hamilton et al., 1993), investors still sought socially responsible assets. This led to the theorisations that ESG investors receive non-financial utility from investing in line with their ethical and social values (Benson and Humphrey, 2008; Renneboog et al., 2008; Renneboog et al., 2011; Pástor et al., 2021b). However, recent survey evidence reveals that the majority of investment managers use ESG data because they perceive a firm's ESG performance to be material to its financial performance as an investment (Amel-Zadeh and Serafeim, 2017). Regardless of whether ESG investing is ethically or financially motivated, investor interest in ESG criteria places expectations on companies to engage in and disclose ESG-related practices.

There is plenty of literature on firm motives for being more ethical, or at least appearing to be so. Evidence shows that many companies engage in CSR out of a genuine desire to be good corporate citizens, even if it is detrimental to shareholder value (Moser and Martin, 2012). Another category of motives, on the other side of return maximisation, is risk reduction. There is vast evidence in favour of firms investing in CSR as a way of mitigating risk, mainly legal and reputational (e.g., Devinney, 2009; Miles and Covin, 2000; Van Duuren et al., 2016; Weber, 2008). In other words, a firm may invest in CSR to shield against liability and controversy, attempting to distract investors and the general public from the company's ethical failings (e.g., Godfrey et al., 2009; Schnietz and Epstein, 2005).

This brings us to another theme around the reliability of CSR scores: to what extent can they be manipulated by management, not only through selective disclosure (Deegan, 2002), but also through the pursuit of 'soft' CSR (Li et al., 2019) with the aim of portraying better corporate citizenship than is true? This is the question of whether firms are involved in 'greenwashing'. Greenwashing is a term that dates back to 1986 when coined by Jay Westerveld in his criticism

of hotels' pretention that their reuse of towels makes them environmentally friendly, despite the hotels' high environmental impact in other critical areas (Pearson, 2010). It has since been the subject of numerous studies questioning the legitimacy of social responsibility claims (e.g., Laufer, 2003) as well as exploring greenwashing's potential downsides at the firm and market levels (e.g., Arouri et al., 2021; Kim and Lyon, 2015; Marquis et al., 2016). A related practice is 'window-dressing', aptly defined by Connors et al. (2017) as "self-serving enhancements undertaken by companies in order to distract from larger truths or underlying problems".

With an incentive to manage their image, companies are more likely to window-dress their CSR scores by engaging in 'symbolic CSR' rather than 'substantive CSR' (Li et al., 2019). In fact, there is evidence that companies which communicate their CSR information with a self-promotional tone, trying to garner consumer trust, see the opposite effect occur and their reputations suffer (Kim, 2019). Mecaj and Bravo (2014) find that a firm invests in low-cost, 'soft' ESG dimensions, such as the adoption of unenforced diversity policies, when it becomes aware of its impending distress. The possibility of 'massaged' CSR scores distorting the CSR-CFP link is explored in our first empirical chapter that focuses on the Subprime and Sovereign Debt Crises in Europe.

A similar natural experiment to the Subprime Crisis explored by Lins et al. (2017) is lent by the COVID-19 crash observed in the early months of 2020. This crisis, driven by a public health concern over the Coronavirus pandemic, is the focus of our second empirical chapter. At the very end of 2019, the World Health Organization (WHO) was notified of pneumonia outbreaks with an unknown cause (UK Security Health Agency, 2021). On January 30, 2020, the WHO declared COVID-19 a 'Public Health Emergency of International Concern', and a global pandemic shortly thereafter on March 11, 2020 (WHO, 2020). However, by that point, this classification was "not a surprise", with the novel virus having been detected in over 100 countries (Roxby, 2020). As of May 2022, the worldwide COVID-19 death toll was over 6.2

million.[3] The COVID-19 pandemic also brought about a recession. From a market perspective, several stock indices lost over 20% of their value during the first quarter of 2020.[4]

The pandemic gave rise to several studies once again considering the responsibility-return relationship in light of an exogenous shock. Interestingly, the findings of this niche CSR literature are persistently heterogenous. Ding et al. (2020) were the first to consider international, cross-firm stock price reactions to the COVID-19 pandemic. Using a sample of over 6,000 firms across 56 economies, they find that firms with higher pre-crash CSR scores suffered milder losses during the first quarter of 2020, identified as the COVID-19 crisis. Similarly, Albuquerque et al. (2020) find that firms with higher ES (Environmental and Social) scores earned an extra daily return of 0.45% during a differently identified crisis peak and US-only sample. They hypothesise that the positive return differential comes through the customer loyalty channel in line with previous work showing that high-ES firms are better able to maintain customer demand (Albuquerque et al., 2019; Servaes and Tamayo, 2013).

On the other hand, a more recent US analysis by Bae et al. (2021) negates these results and finds no evidence that CSR scores possessed return explanatory power during the COVID-19 crash. They challenge earlier work on the subject by suggesting that findings are sensitive to sample selection, regression specification, and recency of CSR data used. The lack of a consensus on this question allows room for further investigation into whether nonfinancial factors relating to stakeholder welfare contributed to COVID-19 stock performance. This is explored in our second empirical chapter.

[3] Worldometer (2022) accessed via Statista; https://www.statista.com/statistics/1093256/novel-coronavirus-2019ncov-deaths-worldwide-by-country.
[4] Il Sole 24 Ore. (2020) accessed via Statista; https://www.statista.com/statistics/1105021/coronavirus-outbreak-stock-market-change.

An element shared between these aforementioned studies is the use of multidimensional ESG scores to measure CSR. These are often obtained from MSCI and/or Refinitiv's ESG database. While both providers pre-calculate an overall ESG score, CSR research often calls for the recalibration of the measure in line with various intuition. For example, it is relatively standard to exclude the Governance Pillar components or keep from them only the dimensions relating to a firm's CSR strategy (e.g., Ding et al., 2020; Lins et al., 2017).

The deconstruction of aggregate scores goes further. In hypothesising that the COVID-19 pandemic has raised concerns over the future of climate actions, Garel and Petit-Romec (2021) consider the responsibility-return relationship strictly from an environmental perspective. As such, they make use of only the environmental score from Refinitiv and find a positive association with performance. The study deconstructs the score further and finds that the association is stronger when the more climate-based indicators, such as reduction in emissions, are isolated and used to measure responsibility. Meanwhile, Shan and Tang (2020) consider the CSR-CFP relationship from a social perspective, focusing on a firm's workforce welfare.[5] Specifically, using a sample of Chinese firms and data from MioTech (a Chinese ESG-data provider), the authors find that higher employee satisfaction gave rise to better COVID-19 stock performance. More centred studies like these shed light on the relationship between various single-dimensional ESG measures and firm performance. While multidimensional scores can give a rounded measure of a firm's overall commitment to CSR, some ESG aspects can be intuitively more relevant than others in certain economic, social, and political conditions.

[5] We cannot be certain that the ESG dimensions from different data providers correspond to the same CSR indicators; however, we assume reasonable parity where relevant. For example, we suggest that the "Employee Relations" dimension from MSCI is comparable to the "Workforce" dimension from Refinitiv and acknowledge that they fall under the 'Social' (S) Pillar of ESG. This allows us to draw some comparisons between findings of different studies where ESG ratings from different providers are used.

As suggested by Steffen (2021), aggregate ESG scores limit price discovery by combining diverse and complex indicators which could individually be more informative. In agreement, Demers et al. (2021) strongly contest that ESG scores were positively associated with COVID-19 performance, suggesting that studies finding such a relationship suffer from incomplete controls. Instead, the authors find that COVID-19 firm performance was immunised by investment in intangibles, measured using accounting entries. However, previous work indicates that investments in social capital and in intangibles are not necessarily discrete, creating overlap between measures for the two (e.g., Shen et al., 2019; Vilanova et al., 2009).

Related to the intangibles of 'human capital' and 'company culture', an area that has garnered attention during and since the COVID-19 pandemic, is employee work flexibility. With widespread workplace closures and limitations on travel and contact, as well as added childcare responsibilities due to school closures, many workers had to adapt to a new working lifestyle. Though estimates vary, statistics indicate that more than half of the workforce in the US and the UK was working from home (WFH) at some point during the pandemic (Bick et al., 2020; Brynjolfsson et al. 2020; Elliott, 2020). This dimension of workforce-related CSR and its importance during the COVID-19 crash is the focus of our second empirical chapter. We contribute to the literature that has found higher industry work flexibility to have been associated with better financial performance (e.g., Alon et al., 2020; Favilukis et al., 2020; Papanikolaou and Schmidt, 2020). We expand on this area in Chapter 4's overview of the relevant literature (Section 4.3).

Our third empirical chapter aims to contribute to the growing literature on investor sentiment towards environmental responsibility. The United Nations (2022) describes climate change as the "defining issue of our time".[6] Bloomberg Intelligence indicate that a major contributor to

[6] UN Global Issues; https://www.un.org/en/global-issues/climate-change.

the growth of ESG assets, which have surpassed a value of USD 35 trillion in 2020, is ESG debt, including green bonds (Bloomberg, 2021). The green bond market reached a cumulative issuance valuation of USD 1 trillion in 2020, after having nearly doubled each year on average since its inception in 2007 (Climate Bonds, 2022). According to Schroders Global Investment Study for 2021, more than half of the 23,950 investors surveyed globally said that sustainability has become more important to them since the pandemic, and more investors are agreeing that major shareholders are themselves responsible for mitigating climate change (Gulliver-Needham, 2021).

This increase in regard for sustainability in finance and investing has been accompanied by diverse academic research on climate change. Among the questions being asked is what role, if any, does the 'greenness' of an asset play in its pricing? Bolton and Kacperczyk (2021) consider this from the perspective of carbon risk, exploring the relationship between a firm's carbon emissions and its stock returns in the US. They find evidence of a carbon premium, indicating that investors are already demanding compensation for exposure to carbon risk. This is in line with other research that links higher firm carbon emissions to greater risk (e.g., Hsu et al., 2021; Ilhan et al., 2020).

There is evidence that green assets have lower betas, and hence command lower returns during normal times (e.g., Choi et al., 2020; Engle et al., 2020). Engle et al. (2020) proceed to show that the higher risk of brown assets can be hedged using investment in green assets, where 'greenness' is determined using environmental ratings. This is on account of green assets earning positive alphas when the climate deteriorates unexpectedly (Choi et al., 2020; Engle et al., 2020). These studies are discussed in more detail in Chapter 5's overview of the related literature (Section 5.3), as this research relates to our focus on an unexpected climate incident – the 2021 summer floods in Europe.

A subset of the climate literature investigates stock market reactions to extreme weather events. Wang and Kutan (2013) focus on the Japanese and US markets and show that significant changes in value during natural disasters are only observed in the insurance sectors. Interestingly, the relationship is different in the two countries: insurance companies lose value in the US while they perform well in Japan. Worthington (2008) looks at the Australian stock market and finds no significant relationship between natural disasters, including floods, earthquakes, and bushfires, on stock market returns. As extreme weather events are on the rise (UNDRR, 2020)[7], it becomes increasingly important to understand their effect on financial markets and the economy more broadly.

We observe that the relationship between greenness and market performance is underexplored in Europe, despite the fact that the Climate Change Performance Index 2022 places three European countries at the top of its rankings (CCPI, 2022). Data from Reuters (2021)[8] also indicates that Europe is the largest market for sustainable funds worldwide. This would strongly suggest that Europe is generally an environmentally conscious region. There is evidence that the effect of an asset's greenness or resilience to climate change on its value is dependent on local sentiment about climate change's legitimacy (Baldauf et al., 2020). As such, our third and final empirical chapter contributes to the environmental responsibility literature by focusing on the effect of climate deterioration on European equity markets.

The review of the literature presented above shows that significant efforts have been made towards clarifying the CSR-CFP relationship, but it remains highly ambiguous. The majority of the work still relies on aggregate ESG ratings as a measure of responsibility and sustainability. However, there is good reason to believe that some CSR dimensions are more

[7] United Nations Office for Disaster Risk Reduction (2020); https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019.

[8] Reuters (2021) accessed via Statista; https://www.statista.com/statistics/1296334/sustainable-funds-asset-size-by-region.

informative than others, at least in certain conditions. In this thesis, we challenge the reliance on only aggregate ESG scores by disentangling them, exploring the relationship between a CSR dimension that is most contemporaneously relevant and firm equity performance during specific windows of time.

Furthermore, we discuss how some dimensions of CSR may be more verifiable and less subject to manipulation than others, which may make them superior to aggregate scores in evaluating sustainability. We also include European firms in our sample to further globalise ESG research which remains US-centric in many areas. Understanding how different regions regard CSR can be highly informative for investment decision making, regulation, and company strategy.

Moreover, in its nature, ESG investing is an evolving practice. As investors become more ESG-conscious, in part due to social and environmental concerns heightening globally, our understanding of ESG investing – its motives, manifestations, pros, and cons – needs to be constantly revised and explored. By investigating very recent sample periods, including the COVID-19 pandemic and the 2021 summer flooding in Europe – the latter of which has not yet been the subject of any academic finance research, to our knowledge – we contribute to the current understanding of the CSR-CFP link, which will set the ground for future work.

3 THE SUBPRIME AND SOVEREIGN DEBT CRISES IN EUROPE: THE ROLE OF SOCIAL RESPONSIBILITY

3.1 Abstract

In the Subprime Crisis of 2008/9, European firms with high CSR scores had stock returns 1.8 to 2.6 percentage points higher than those with lower scores. This is believed to be the effect of social capital – an asset which becomes incrementally valuable during crises of trust. However, during the subsequent Sovereign Debt Crisis in 2011, we find the opposite to hold true. European firms with high CSR scores had stock returns which were roughly 1.6 percentage points lower than the returns of low-CSR firms. We argue that the drastic change in the market response to CSR scores may be the result of the scores becoming a proxy for ESG controversies associated with legal disputes and fines. We find that using a new measure of social responsibility based on firms' efforts to avoid ESG controversies eliminates the return penalty previously observed with raw CSR scores. We conclude that European market participants have not ceased to value social capital in their investment decision-making following the Subprime Crisis.

3.2 Introduction

Since the mid-2000s, CSR has been strongly associated with social capital and societal trust. When a company invests in CSR, it increases its social capital. This gives rise to the public establishing a feeling of connectedness with the company and thereby developing 'trust' in it (Sacconi and D. Antoni, 2011). Intuitively, trust is linked to times of financial crisis because it is during those times that trust in the financial system dwindles. As such, one can posit that companies which hold onto their investors' trust have a scarce and valuable resource. To empirically test this, Lins et al. (2017) conduct a study in which they investigate the relationship between CSR, as a proxy for social capital, and Subprime crisis-period returns in the US. They find that high-CSR firms, which have built stronger trust with investors, earned returns four to seven percentage points higher than low-CSR firm returns during the crisis window.

We test whether this finding is robust in markets other than the US, and during crisis periods other than the Subprime Crisis, by looking at the European market during both the Subprime Crisis and the Sovereign Debt Crisis of 2011. We use the Refinitiv ESG database to investigate the relationship between CSR and crisis performance. We find that, in Europe, high-CSR companies perform better than low-CSR companies during the Subprime Crisis by up to 2.6 percentage points, on average, which is consistent with the findings of Lins et al. (2017). However, we show non-linearity in this association whereby the best performance is earned not by companies in the top two CSR quartiles, but by those in the best and second-worst ones. We point to previous literature to explain this phenomenon.

By contrast, we detect a strong negative correlation between CSR scores and crisis-period returns during the Sovereign Debt Crisis. This is robust to the use of different crisis measurement windows. We explain this unexpected outcome by showing that there is a positive correlation between CSR scores and firm involvement in sustainability-related controversy which could lead to legal disputes and fines. We find that controversies were much stronger in

the years leading up to the Sovereign Debt Crisis than they were prior to the 2008 crisis. We also find that the correlation between CSR scores and controversies increased in the later crisis. Given the positive association we find between a firm's controversy avoidance and its financial performance around the Sovereign Debt Crisis, we suggest that controversy avoidance might have become superior to raw CSR scores in signalling responsibility.

This paper contributes to the literature in three main ways. First, we test the robustness of previous findings that link CSR to positive stock returns in crisis periods and show that such a relationship is highly dependent on the sample period and geographic region. We also show that the information content of CSR scores may be changing over time, altering the relationship between the scores and stock market responses. Second, we employ an underutilised CSR indicator, which we term "Controversy Avoidance". Previous work suggests that it is difficult to provide a consensus on the theoretical meaning of CSR (Dahlsrud, 2008), so wide disagreement continues to exist on the validity of its various indicators. Our results show that using a measure of sustainability which captures a firm's controversy avoidance has merit, at least in explaining more recent crisis-period performance in Europe. Third, we contribute to the literature that investigates firm motives in pursing CSR initiatives (e.g., Chatjuthamard-Kitsabunnarat et al. 2014; Hoepner and Schopohl, 2020; Weber, 2008; Withisuphakorn and Jiraporn, 2018). We do so by bringing to light a strong correlation between firm CSR scores, which are subject to influence by direct corporate action, and year-ahead controversy. We suggest that firms act on the incentive to window-dress their CSR scores in anticipation of impending negative publicity.

The rest of this chapter is structured as follows. Section 3.3 summarises the motivation for this research by briefly reviewing relevant literature and describes the main hypothesis tested. Section 3.4 describes the European Subprime and Sovereign Debt Crisis samples employed in our analysis. Section 3.5 presents our results using CSR scores as the measure of responsibility.

Section 3.6 explores the use of controversy avoidance, our alternative CSR measure, and Section 3.7 concludes.

3.3 Related Literature and Motivation of Study

Our analysis in this empirical chapter builds on Lins et al.'s (2017) work which show that firms with higher CSR scores had better stock performance during the Subprime Crisis of 2008. The authors underpin their hypothesis by conceptually linking 'trust', 'social capital' and 'CSR'. Both trust and social capital, often used synonymously in the economics literature, revolve around the theme of cooperation in producing positive outcomes (e.g., Fukuyama, 1995; Lopez-de-Silanes et al., 1997). In practice, a firm investing in social capital by, for example, making donations, positively contributes to the community (Guiso et al., 2011; Scrivens and Smith, 2013). This, in turn, fosters trust in said firm.

The literature linking trust and investment is extensive. Zak and Knack (2001) find that investment is higher in high-trust regions. This is, in part, because shareholders can reduce monitoring costs when they trust their agents (Knack and Keefer, 1997). This is corroborated by Balloch et al. (2015) who find that higher trust in the stock market increases participation, as in Georgarakos and Pasini (2011), as well as the size of participant shareholdings. Guiso et al. (2008) show that more trusting individuals are more likely to become shareholders, and that the perception of how trustworthy a company is plays a part in whether its shares are bought. Furthermore, if a firm is trustworthy, its stakeholders are more willing to cooperate with it; for example, employees would work more diligently (Guiso et al., 2015) and customers would be more loyal (Servaes and Tamayo, 2013). Knowing this, shareholders may be inclined to pay a premium for trustworthy firms in expectation of their operations being more resilient during periods of low market confidence (Lins et al., 2017).

One way to measure societal trust in a firm – i.e., its social capital – is through its CSR score. This is on account of the closeness in definition for both concepts (World Business Council for Sustainable Development, 2000). CSR's role in building social capital is also evidenced in the literature (e.g., Bénabou and Tirole, 2010; Eccles et al., 2014). One can put forth that firms with better CSR scores are considered more trustworthy by investors, and this characteristic ought to become more valuable when the market suffers a negative shock to trust, as it did during the Subprime Crisis. Lins et al. (2017) build this hypothesis and find robust evidence in favour of it.

An interesting question is whether their findings are the same, similar, or reversed in another large market during the same crisis and a later crisis a few years thereafter. We question whether the European Sovereign Debt Crisis which ensued in 2011 was also a crisis of trust, with the two crises inextricably linked. We posit that the substantial government bond downgrades and subsequent bailouts, as experienced by Greece and Portugal, for example, further undermined investor confidence in European capital markets. This is in line with perceptions that Europe experienced stunted subprime recovery due to its chaotic political response and lack of crisis management preparedness (Lane, 2012). Therefore, we look at the relationship between CSR and performance during both periods, comparing them not only to each other but also to previous findings in the US market. By exploring the same hypothesis within different samples, both cross-sectional (US versus European) and temporal (the 2008 Subprime Crisis versus the 2011 Sovereign Debt Crisis in Europe), we offer new insights on the CSR-CFP link, which remains highly debated.

The link between trust and performance during crises could have meaningful implications for regulators. Regulators may be incentivised to encourage higher CSR engagement in order to increase societal trust, hoping it would translate to more stable markets and improved economic conditions (Guiso et al., 2004, 2008; Leonardi, 2001). Considering that financial crises can

arise from investors', depositors', and consumers' fear of loss, the higher the trust the public has in financial markets, the less pronounced the negative effects of an economic downturn could be.

We also contribute to the literature on robust CSR measurement by shedding light on an indicator which does not rely on companies' self-reporting. Rather, our measure based on controversy avoidance is media-determined, relying on whether the firm had been involved in publicised controversy issues. The media can influence the public's image of a firm and thereby indirectly impact the firm's performance, in line with stakeholder theory (Henriques and Sadorsky, 1999). While companies are eager to disclose positive information about their CSR practices, information about a firm's social irresponsibility is more likely to be reported through the media (Berkan et al., 2021; Dyck et al., 2010). Adverse media coverage can increase a firm's reputational risk and make investors less certain about the firm's financial prospects (Dyck and Zingales, 2002). This reputational damage can be an indirect channel through which a firm's financial performance is negatively affected (Graham et al., 2008; Guiso et al., 2006). Social irresponsibility can also directly impact financial performance due to increased litigation costs (Murphy and Coombes, 2009). The evidence linking social irresponsibility to unfavourable firm outcomes explains why firms may wish to distract from their ESG-related controversies, possibly through the pursuit of CSR initiatives (Kang et al., 2016; Yue et al., 2022). We argue that gauging responsibility through a media-based measure of how well a firm has avoided controversy is less manipulable by management, and hence has potential to clarify the true CSR-CFP relationship, at least during crisis periods.

Mecaj and Bravo (2014) find that a firm takes insubstantial, low-cost action, such as the adoption of unenforced diversity policies, when it becomes aware of its impending distress. This suggests that firms can not only window-dress their image in reaction to bad news (e.g., Kotchen and Moon, 2012) but they can also do so proactively in anticipation of future trouble.

In other words, companies can foresee impending controversy and act on this motive to remedy their CSR image ahead of time. CSR being used in this way is often described as an 'insurance mechanism' for protection against future corporate social irresponsibility (Kang et al., 2016). This is underpinned by the supposition that a firm investing in CSR effectively builds reputational goodwill that can safeguard the firm's financial performance when it is subject to negative attention on account of social or environmental failings (Godfrey et al., 2009; Jones et al., 2000; Minor and Morgan, 2011; Schnietz and Epstein, 2005). However, evidence suggests that corporate social responsibility and irresponsibility are not additive (Kang et al., 2016; Oikonomou et al., 2014); as such, a firm cannot compensate for ethical wrongdoing simply by investing in CSR. Building on this, we put forth that investing in CSR as an insurance mechanism may in fact become detrimental to companies in the long run if stakeholders come to believe that the most CSR-intensive firms are possibly also the most controversial. We contribute some support to this theory on firm motives for pursuing CSR by detecting that CSR scores, which can be directly influenced by corporate action, signal year-ahead controversy.

3.4 Data and Methodology

3.4.1 Sample Construction

The study is conducted using a cross-sectional regression of crisis-period holding returns on a measure of CSR and a set of control variables. We follow the methodology of Lins et al. (2017) as closely as possible. The starting point was finding all those European-domiciled companies which have Refinitiv ESG scores for the required years. In line with previous work, we eliminate financial companies from our samples since they may have received substantial support from governments during the crises. We also exclude micro-cap stocks with a market capitalisation below EUR 250 million as of the year-end preceding each crisis observation

period because they have lower liquidity and higher price pressure, particularly during times of financial distress.

As in Lins et al. (2017), we use a Subprime Crisis observation period beginning in August 2008 and lasting until March 2009. We observe the Sovereign Debt Crisis to be from April 2011 to September 2011. Identifying a window for the peak of the Sovereign Debt Crisis is challenging because different European countries were most severely impacted at different times. Some studies define the crisis period to be over two years long, encompassing several other major events. Lane (2012) suggests a timeline which starts when Greece received a EUR 110 billion bailout and ends when most ten-year government bond yields began to show signs of sustained recovery. This would be from May 2010 until June 2012, when a European Council meeting led to the agreement of bank stability mechanisms (Lane, 2012). However, this period is quite long to be considered a 'peak', with European financial markets experiencing several upward and downward swings throughout these many months. For this reason, we isolate the crisis peak to the period of worst equity performance within this timeframe of a value-weighted composite index of the GIIPS countries' (Greece, Ireland, Italy, Portugal and Spain) main market indices. We create this and observe which combination of entry and exit months gives the worst holding period return. Since many of the companies in our sample are from the United Kingdom, we repeat this process with the FTSE100 index, and both iterations show the worst period to be the second and third quarters of 2011, starting in April 2011 and ending in September 2011, which we select for our regression.

We use the 2006 and 2009 Refinitiv ESG scores for the Subprime and Sovereign Debt samples, respectively. As explained in Lins et al. (2017), the reason the scores are lagged by two years is because those lagged by only one could have been unrepresentatively high following companies' revision to their CSR mandates in anticipation of the crises. To avoid any reverse-causality concerns, the authors used scores from 2006 instead of 2007, believing the former to

be completely free of influence from the Subprime Crisis. For our European analyses, we follow the same approach. To address concerns that the 2009 scores we use for our later crisis are still affected by the Subprime Crisis, we repeat that part of our analysis using 2010 scores; our results do not change in any notable way.

Unlike with MSCI ESG scores, Refinitiv scores are not calculated explicitly on the basis of 'strengths' and 'concerns', i.e., respectively, the positive and negative ESG traits of a firm. They are percentile rank scores which are benchmarked against industry groups and calculated on the basis of ten main themes. Under the Environmental Pillar, the themes are Resource Use, Emissions, Innovation. Under the Governance Pillar, the themes are Management, Shareholders, CSR Strategy. Under the Social Pillar, the themes are Workforce, Human Rights, Community and Product Responsibility (Refinitiv, 2021). In their paper, Lins et al. (2017) exclude all Governance indicators, as is common in the CSR literature, on account of Governance not being a direct part of a company's CSR remit. Similarly, we average only the Environment and Social Pillars of Refinitiv's ESG scores for our *CSR Score* measure.[9] Finally, to bring closer the interpretation of our *CSR Score* variable to typical CSR scores in previous studies, we scale the variable to range between -1 and +1.[10]

The cross-sectional regression model we employ is as follows:

$$\text{Crisis-Period Holding Return}_i = \alpha + \beta_1 \textit{CSR Score}_i + + \beta_2 \textit{Ln(Market Cap)}_i + \beta_3 \textit{Long-Term Debt}_i + \beta_4 \textit{Short-Term Debt}_i + \beta_5 \textit{Profitability}_i + \beta_6 \textit{Cash Holdings}_i + \beta_7 \textit{Book-to-Market}_i +$$

[9] *CSR Score* as a variable is only used in this chapter. Within this chapter, it refers to the average of the E (Environment) and S (Social) scores. This is for comparability with Lins et al. (2017) and to differentiate it from the controversies measure to be used later in this chapter. Outside of this chapter, a similar score based on only the E and S scores will be referred to as the "ES-score".

[10] This is done by subtracting 50 from the percentile score, equating a score of 50/100 to 0, then dividing by 100.

$\beta_8 Negative\ B/M_i + \beta_9 Momentum_i + \beta_{10} Idiosyncratic\ Risk_i + Four\text{-}Factor\ Loadings + Sector\ Dummies + Country\ Dummies + e_i$ *(Equation 3.1)*

For both crises, Crisis-Period Holding Return is calculated twice: once as a raw return, then as an abnormal return. *Crisis-Period Raw Return* is measured as the holding period return throughout the observation period. *Crisis-Period Abnormal Return* is the difference between that holding period return and the expected holding period return as predicted by the Capital Asset Pricing Model (CAPM), which is run using five years of monthly observations prior to the respective start date of each crisis, using Kenneth French's website for European asset pricing factors. In addition to *CSR Score*, the key independent variable, the other regressors control for firm characteristics, risk, and previous performance. The accounting variables capturing firm characteristics are *Long-Term Debt*, *Short-Term Debt*, *Cash Holdings* and *Profitability*, all of which are included as a ratio to Total Assets. Furthermore, equity *Book-to-Market* and a dummy for *Negative B/M* firms are included, where the latter corrects for the behaviour exhibited by negative book-to-market firms which are likely to be in distress and hence act more like high book-to-market firms (Fama and French, 1992). The accounting variables are measured as closely as possible to the year-end preceding the crisis observation period. This would be the last quarter of 2007 for the Subprime Crisis and last quarter of 2010 for the Sovereign Debt Crisis. Furthermore, *Momentum* is included to control for the firm's performance leading up to the crisis and is measured as the firm's raw return over the 12 months preceding the crisis observation period. Finally, due to evidence that a firm's stock price volatility impacts its returns (Goyal and Santa-Clara, 2003), *Idiosyncratic Risk*, measured as the residual variance from the aforementioned CAPM regressions, is included. Carhart (1997) four-factor loadings based on regressions of monthly data for five years preceding each crisis[11],

[11] As in Lins et al. (2017), companies with fewer than 12 months of observations for the market model and four-factor regressions are excluded from the sample.

sector dummies, and country dummies are also used to control for further effects. All variables are winsorised at the 1st and 99th percentiles to mitigate the effect of outliers, and White (1980) heteroscedasticity-consistent standard errors are used. All financial and accounting data is obtained from Refinitiv via Thomson Reuters Eikon (unless otherwise stated).

All European companies with this data available are included in their respective sample(s). Since the number of companies rated by Refinitiv increases over time, this naturally meant that our largest-possible Subprime Crisis sample was smaller than its Sovereign Debt Crisis counterpart, containing 487 and 598 observations, respectively. Performing the analysis on a common sample of 486 firms which have enough data coverage for both crises does not change the results we had when using the largest-possible respective samples in any notable way. Consequently, to make the results from the tests of both crises as comparable as possible, we only report the statistics and results of the common sample whose statistics are described below.

3.4.2 Descriptive Statistics

Table 3.1 shows the summary statistics and correlation matrices during both crisis observation periods for our main variables. In Panels A and B, subpanels (i) for the Subprime Crisis show the respective statistics for the Lins et al. (2017) US sample in the column before the vertical line, included for accessible comparison. While there are some minor differences between our Subprime Crisis sample statistics and Lins et al.'s (2017), the majority of the trends and figures are similar. Looking at *Crisis-Period Raw Return*, both US companies and European companies made holding period losses of approximately 30% to 40%, which confirms that investors were concerned about companies' survival and acted on their fears. As for the Sovereign Debt Crisis sample, it can be seen that the *Crisis-Period Raw Return* is not as negative as it was for the Subprime Crisis. However, considering an average loss of 20.7% was suffered over a shorter, six-month period, its severity may be comparable to the 2008 crisis in

Europe. We also note that, on average, *CSR Score* more than doubled from 2006 to 2009, and its distribution is positively skewed. This may be intuitive considering stakeholders' increased attention to non-financial performance in more recent years, which could have led to firms becoming more engaged in CSR and/or CSR reporting. We also explore an alternative explanation for why the CSR scores seem inflated in the years following the Subprime Crisis and report our findings in Section 3.6.

Figure 3.1 summarises the key difference we find in the relationship between returns and CSR from one crisis to another. Each bar on the horizontal axis represents a European company from our sample, and its return is represented on the left vertical axis. These bars are sorted in ascending order, from worst to best return, for each crisis. It is clear from the return plots that the two crises manifested similarly, with the Sovereign Debt Crisis being slightly milder than its predecessor. We also show the trendlines of each crisis's returns against respective CSR scores. These constitute 'lines of best fit', linearly approximating the relationship between firm returns and CSR scores during each crisis. The Subprime Crisis trendline, labelled Linear (CSR 2006), has a large, positive gradient while the Sovereign Debt trendline, labelled Linear (CSR 2009), is far less steep. This would imply that the positive association between returns and CSR scores weakened after the first crisis. The figure reiterates that the average *CSR Score* during 2009 was higher than in 2006, implying that companies have been engaging in CSR initiatives more actively following the Subprime Crisis. It is worth questioning what the company's incentives were in doing so if social responsibility was indeed no longer valued by the market. To understand these changing trends better, Sections 3.5 and 3.6 investigate these relationships in a regression framework with the employment of control variables and the discussion of alternative explanations.

It is important to bear in mind that our analysis incorporates 19 European countries as opposed to a single market like the US. It is therefore interesting to see the variation in performance

from one country to another, which reflects the fact that some countries rode out the 2011 storm better than others. We report a 'by country' summary of our sample in Table 3.2, Panel A and add a 'by sector' summary for reference in Panel B. The United Kingdom contributes a large portion of firms to our sample due to higher Refinitiv coverage in Britain. However, the average UK values for both crisis raw returns and CSR scores are not out-of-line with other countries'. Therefore, the sizeable presence of British companies should not steer our results in a way that is unrepresentative of Europe as a whole.

In line with expectations, the worst performing firms during the Sovereign Debt Crisis are Greek companies with an average return of -49%. Perhaps more surprisingly, Germany's subprime return is highly negative at -45%. This is not far off from the DAX's buy-and-hold loss of 38% over the same observation window. In our sample, seven German companies lost over 70% of their value during the Subprime Crisis, explaining the discrepancy between our value and the DAX's. Although data coverage is not uniform across countries, we have sufficient diversity in our sample to have good representation of Europe as a region. We control for country-specific effects by including country dummies in our regression models.

The correlations in Table 3.1, Panel B reveal the first notable finding of this paper. While our Subprime Crisis sample exhibits very similar patterns to those in Lins et al.'s (2017) paper for the US, with positive correlations of 0.14 and 0.04 between *CSR Score* and crisis raw and abnormal returns, respectively, the Sovereign Debt Crisis figures tell a different story. While the correlation between *CSR Score* and raw returns is mildly positive at 0.04, the correlation with abnormal returns is negative. This is an indication that the relationship between CSR and performance might have changed over time. We explore these relationships further in the next sections.

Figure 3.1. Comparison of the Relationship between Crisis-Period Returns and CSR Scores across Crises

The common sample consists of 486 firms with CSR data available from the Refinitiv ESG database as of year-ends 2006 and 2009 and returns available during the period August 2008 to March 2009 and April 2011 to September 2011. *Crisis-Period Raw Return* is the raw return computed over the periods August 2008 to March 2009 and April 2011 to September 2011 for the Subprime and Sovereign Debt Crisis, respectively. Each bar on the horizontal axis in this figure represents a European company from our common sample. Firm crisis-period raw returns are represented on the left vertical axis. For each crisis independently, the bars are sorted in ascending order, from worst to best return. The values of *CSR Score* range from -1 to +1 and are the rescaled average of Refinitiv's Environment and Social Pillar Percentile Scores. *CSR Score* is measured at the end of 2006 and 2009 for the Subprime and Sovereign Debt Crisis, respectively. In this figure, *CSR Score* is plotted on the right vertical axis, which is bounded at 0 and 0.3 surrounding average values, with the individual *CSR Score* data points hidden for chart readability. We only show the trendlines of firm returns versus CSR scores for each crisis: Linear (CSR 2006) is the trendline for the Subprime Crisis, and Linear (CSR 2009) is the trendline for the Sovereign Debt Crisis.

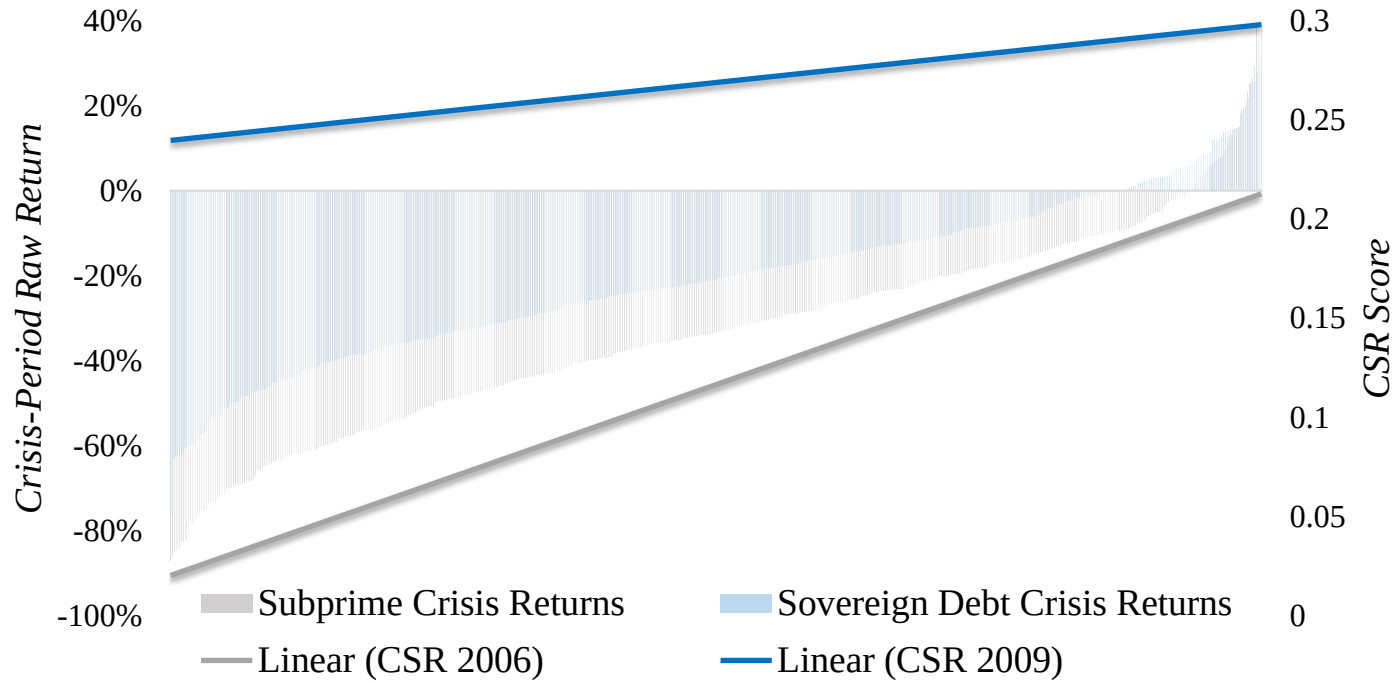


Table 3.1. Descriptive Statistics: Subprime and Sovereign Debt Crises and CSR

The common sample consists of 486 firms with CSR data available from the Refinitiv ESG database as of year-ends 2006 and 2009 and returns available during the period August 2008 to March 2009 and April 2011 to September 2011. *CSR Score* ranges from -1 to +1 and is the rescaled average of Refinitiv's Environment and Social Pillar Percentile Scores. CSR ratings are measured at the end of 2006 and 2009 for the Subprime and Sovereign Debt Crisis, respectively. *Crisis-Period Raw Return* is the raw return computed over the periods August 2008 to March 2009 and April 2011 to September 2011 for the Subprime and Sovereign Debt Crisis, respectively. *Crisis-Period Abn. Return* is the market model-adjusted return over the above periods, with market model parameters computed over the five-year period ending in July 2008 and March 2011 for each crisis respectively, using the K. French European value-weighted index as the market proxy. Accounting data is based on the last quarter ending at or before the ends of 2007 and 2010 for each crisis respectively and is obtained from Thomson Reuters Eikon. *Market Capitalisation* is in millions of euros. *Long-Term Debt* is computed as long-term debt divided by total assets. *Short-Term Debt* is computed as debt in current liabilities divided by total assets. *Cash Holdings* is computed as cash and marketable securities divided by total assets. *Profitability* is computed as operating income divided by total assets. *Book-to-Market* is computed as book value of equity divided by market value of equity. *Negative B/M* is a dummy variable set to one when the book-to-market ratio is negative and zero otherwise. *Momentum* is the raw return over the period August 2007 to July 2008 and April 2010 to March 2011 for each crisis, respectively. *Idiosyncratic Risk* is computed as the residual variance from the market model estimated over the five-year period ending in July 2008 and March 2011, respectively, using monthly data. Financial firms and micro-cap stocks, defined as firms with a market capitalisation below EUR 250 million as of year-end 2007 and 2010 for each crisis respectively, are removed from the sample. All financial data is obtained from Refinitiv via Thomson Reuters Eikon (unless otherwise stated). Control variables and returns are winsorised at the 1st and 99th percentiles. For the Subprime Crisis subpanels (i), we include relevant results from Lins et al. (2017) to facilitate comparison. In Panel A, the columns headed "LST '17" correspond to the first two value columns in Lins et al. (2017) Table I, Panel A, showing the summary statistics for that study's US sample. The number of observations in their sample is 1,673. The *Market Capitalisation* values listed in these two columns below are the EUR-equivalent to the USD values in Lins et al. (2017). In Panel B, the column headed "*LST '17 CSR*" corresponds to the first value column in Lins et al. (2017) Table I, Panel B.

Panel A: Summary Statistics, *N* = 486

(i) Subprime Crisis: 8-month period from August 2008 until March 2009

	LST '17		Mean	*SD*	25th perc.	Median	75th perc.
	Mean	*SD*					
CSR Score	-0.165	0.381	0.116	0.349	-0.167	0.081	0.403
Crisis-Period Raw Return	-0.391	0.284	-0.334	0.238	-0.493	-0.333	-0.179
Crisis-Period Abn. Return	0.116	0.592	0.027	0.575	-0.101	0.037	0.172
Market Capitalisation (€m)	4741	16397	17741	29698	2205	6360	18280
Long-Term Debt	0.198	0.193	0.197	0.146	0.082	0.177	0.292

Short-Term Debt	0.029	0.055	0.057	0.060	0.012	0.040	0.083
Cash Holdings	0.172	0.199	0.095	0.092	0.032	0.065	0.123
Profitability	0.033	0.034	0.113	0.080	0.063	0.097	0.148
Book-to-Market	0.430	0.295	0.537	0.739	0.132	0.291	0.598
Negative B/M	0.000	0.155	0.004	0.064	0.000	0.000	0.000
Momentum	-0.082	0.370	-0.252	0.257	-0.442	-0.275	-0.093
Idiosyncratic Risk	0.011	0.010	0.005	0.004	0.003	0.004	0.006

(ii) Sovereign Debt Crisis: 6-month period from April 2011 until September 2011

	Mean	*SD (Std Dev)*	25th perc.	Median	75th perc.
CSR Score	0.269	0.367	0.019	0.322	0.558
Crisis-Period Raw Return	-0.207	0.184	-0.337	-0.208	-0.081
Crisis-Period Abn. Return	-0.043	0.164	-0.146	-0.043	0.065
Market Capitalisation (€m)	17711	29675	2208	6347	18268
Long-Term Debt	0.202	0.141	0.107	0.187	0.290
Short-Term Debt	0.045	0.049	0.010	0.033	0.060
Cash Holdings	0.104	0.090	0.039	0.076	0.144
Profitability	0.092	0.068	0.050	0.076	0.121
Book-to-Market	0.595	0.807	0.134	0.322	0.689
Negative B/M	0.014	0.119	0.000	0.000	0.000
Momentum	0.125	0.237	-0.026	0.107	0.248
Idiosyncratic Risk	0.008	0.007	0.004	0.006	0.010

(Continued…)

Panel B: Correlation Matrix, *N* = 486

(i) Subprime Crisis: 8-month period from August 2008 until March 2009

	LST '17 CSR	*CSR Score*	*C. Raw Return*	*C. Abn. Return*	*Ln (Mkt Cap)*	*LT Debt*	*ST Debt*	*Cash Hold.*	*Profit.*	*B/M*	*Neg. B/M*	*Mom.*
C. Raw Return	0.11	0.14										
C. Abn. Return	0.08	0.05	0.37									
Ln (Mkt Cap)	0.20	0.39	0.15	0.02								
LT Debt	-0.07	-0.01	-0.05	-0.08	-0.11							
ST Debt	0.06	0.06	-0.07	-0.05	0.04	-0.03						
Cash Holdings	0.06	-0.01	0.05	0.06	0.05	-0.39	-0.12					
Profitability	0.05	-0.04	0.16	0.11	0.08	-0.34	-0.11	0.27				
Book-to-Market	-0.09	-0.01	-0.26	-0.05	-0.19	0.05	0.05	-0.06	-0.16			
Negative B/M	-0.02	-0.07	0.04	-0.01	-0.03	0.07	-0.01	-0.01	-0.05	-0.05		
Momentum	-0.08	0.08	-0.01	-0.05	0.22	-0.10	-0.05	0.12	0.13	-0.10	-0.04	
Idiosync. Risk	-0.12	-0.25	-0.13	-0.12	-0.26	-0.04	-0.11	0.13	0.02	0.06	0.00	-0.20

(ii) Sovereign Debt Crisis: 6-month period from April 2011 until September 2011

	CSR Score	*Crisis Raw Return*	*Crisis Abn. Return*	*Ln (Mkt Cap)*	*LT Debt*	*ST Debt*	*Cash Hold.*	*Profit.*	*B/M*	*Neg. B/M*	*Mom.*
Crisis Raw Return	0.04										
Crisis Abn. Return	-0.05	0.89									
Ln (Mkt Cap)	0.36	0.25	0.15								
Long-Term Debt	0.10	-0.01	-0.01	-0.06							
Short-Term Debt	0.08	-0.14	-0.14	-0.01	0.09						
Cash Holdings	-0.09	0.07	0.08	0.06	-0.37	-0.15					
Profitability	-0.07	0.33	0.23	0.15	-0.28	-0.19	0.29				
Book-to-Market	0.03	-0.36	-0.30	-0.18	0.04	0.14	-0.11	-0.28			
Negative B/M	-0.03	0.06	0.05	-0.04	0.14	-0.03	0.00	-0.04	-0.09		
Momentum	-0.16	0.05	0.18	-0.01	-0.12	-0.07	0.09	0.21	-0.17	-0.09	
Idiosyncratic Risk	-0.30	-0.27	0.03	-0.26	0.07	0.00	-0.01	-0.22	0.06	-0.01	0.07

Table 3.2. Country and Sector Profiles

The common sample consists of 486 firms with CSR data available from the Thomson Reuters ESG database as of year-ends 2006 and 2009 and returns available during the period August 2008 to March 2009 and April 2011 to September 2011. CSR ranges from -1 to +1 and is the rescaled average of Refinitiv's Environment and Social Pillar Scores. CSR ratings are measured at the end of 2006 and 2009 for the Subprime and Sovereign Debt Crisis, respectively. *Crisis Raw Return* is the raw return computed over the periods August 2008 to March 2009 and April 2011 to September 2011 for the Subprime and Sovereign Debt Crisis, respectively. Country is determined by country of headquarters as listed on Thomson Reuters Eikon. Sector is the Global Industry Classification Standard (GICS) sector, as per the taxonomy developed by MSCI and Standard & Poor's in 1999 for use by the global financial community, listed on Thomson Reuters Eikon.

Panel A: Descriptive Statistics by Country

		Average by Country			
Country	% of Firms	Subprime 8-Month *Crisis Raw Return*	Sovereign Debt 6-Month *Crisis Raw Return*	2006 *CSR Score* (for Subprime Crisis)	2009 *CSR Score* (for Sovereign Debt Crisis)
United Kingdom	34%	-0.28	-0.13	0.10	0.18
France	11%	-0.32	-0.23	0.31	0.47
Germany	9%	-0.45	-0.23	0.10	0.32
Sweden	7%	-0.28	-0.24	0.13	0.37
Switzerland	7%	-0.36	-0.23	0.09	0.17
Spain	5%	-0.37	-0.24	0.24	0.47
Finland	4%	-0.43	-0.37	0.04	0.20
Netherlands	4%	-0.30	-0.28	0.17	0.41
Republic of Ireland	3%	-0.30	-0.10	0.12	0.17
Belgium	3%	-0.34	-0.23	-0.17	0.22
Denmark	2%	-0.40	-0.24	-0.11	0.07
Italy	2%	-0.33	-0.28	0.14	0.48
Norway	2%	-0.46	-0.26	0.09	0.28
Austria	2%	-0.49	-0.27	-0.05	0.09
Greece	1%	-0.32	-0.45	0.01	-0.02
Portugal	1%	-0.28	-0.19	-0.06	0.43
Luxembourg	0.6%	-0.48	-0.32	0.22	0.39
Jersey	0.4%	-0.41	-0.19	0.07	-0.36
Russia	0.4%	-0.44	-0.26	0.41	0.28

(Continued…)

Panel B: Descriptive Statistics by Sector					
		Average by Sector			
Sector	% of Firms	Subprime 8-Month *Crisis Raw Return*	Sovereign Debt 6-Month *Crisis Raw Return*	2006 *CSR Score* (for Subprime Crisis)	2009 *CSR Score* (for Sovereign Debt Crisis)
Industrials	28%	-0.37	-0.23	0.09	0.19
Consumer Discretionary	13%	-0.25	-0.17	0.14	0.28
Materials	10%	-0.41	-0.29	0.08	0.29
Communication Services	9%	-0.32	-0.22	0.15	0.31
Consumer Staples	8%	-0.19	-0.10	0.16	0.36
Health Care	7%	-0.25	-0.14	0.08	0.22
Energy	7%	-0.40	-0.30	0.14	0.31
Real Estate	7%	-0.41	-0.19	0.07	0.23
Utilities	5%	-0.35	-0.15	0.23	0.49
Information Technology	5%	-0.35	-0.22	0.12	0.22

3.5 Crisis-Period Returns

3.5.1 The Subprime Crisis: Regional differences and the danger of mediocrity

The main regression results for the Subprime and Sovereign Debt Crises are displayed in Tables 3.3 and 3.4, respectively. In each of the tables, Panel A shows how a continuous *CSR Score* impacts crisis returns, both raw and abnormal. It can be seen that our results for the Subprime Crisis are consistent with Lins et al.'s (2017) whereby *CSR Score* is positively and significantly associated with crisis-period returns under all models. When including firm characteristics (Columns 3 and 4), the coefficient of the *CSR Score* is significant at the 10%

level.[12] This means that one standard deviation increase in the *CSR Score* (0.35) is associated with an economically significant 1.8% increase in raw and abnormal crisis-period returns. To put things into perspective, one standard deviation in *Profitability* (0.088), an established contributor to performance, is associated with a 2.9% increase in raw returns, as per Column (3). This means that the contribution from the *CSR Score* is more than half that of *Profitability*, further emphasizing the former's economic significance.

The second analysis presented in Panel B is done using dummy variables for different *CSR Score* quartiles. If a company has a CSR score which falls in the second quartile of all CSR scores in that sample, it will have a *CSR Score Q2* value of one, and zero otherwise, and so on. The coefficients on *CSR Score Q2*, *CSR Score Q3* and *CSR Score Q4* measure the incremental effect of a company being in the second-worst, second-best, and best quartile, respectively. Our results for the Subprime Crisis in European markets diverge in some ways from those in the US. As shown in Panel B, Column (5)[13], Lins et al. (2017) have a monotonically increasing CSR-quartile coefficient when going from one CSR quartile to a higher one, and a positive and significant coefficient for the best quartile. This implies that, all else equal, a company with a higher CSR score achieves higher abnormal crisis-period returns on average after controlling for firm characteristics.

However, our results indicate a more complex effect of CSR on stock returns. While all the *CSR Score*-quartile coefficients are positive, and the largest in each test is on *CSR Score Q4*, which is always significant, the second largest is on *CSR Score Q2*, not *CSR Score Q3*. This suggests non-linearity in the relationship between CSR and performance. The coefficients on *CSR Score Q2* are also significant in all but one specification while those on *CSR Score Q3*

[12] Our significance levels are not as high as Lins et al.'s (2017). This may be due to the fact that we analyse European firms rather than US ones and have a smaller sample size.
[13] This corresponds to Lins et al.'s (2017) Table II, Panel B, Column (4).

only reach significance once and at the 10%-level in Column (2). The implication is that being in the best CSR quartile, *CSR Score Q4*, gives European firms the biggest advantage in terms of financial performance. However, our findings would suggest that a company unable to reach that top quartile is better off in the second-worst quartile, *CSR Score Q2*, as opposed to the second-best quartile, *CSR Score Q3*, since the former's coefficient is always greater than the latter's. One way to justify this finding is to think of public perception. Companies in *CSR Score Q2* reap the benefits of just avoiding the worst, most scrutinised quartile without spending too much on building social capital. Meanwhile, those in *CSR Score Q4* are clearly 'leaders', thereby resonating with and attracting CSR-sensitive investors. Companies in *CSR Score Q3* are likely to incur higher CSR costs than those in *CSR Score Q2*, but they still fall short of reaching the 'elite' *CSR Score Q4* tranche which receives the positive attention.

Support for this reasoning is found in the literature which documents how the initial pursuit of CSR initiatives has the greatest impact on a company's performance as compared to further, more extensive CSR investments (Barnett and Salomon, 2012; Lankoski, 2009; Meijer and Schuyt, 2005; Oikonomou et al., 2012; Wood and Jones, 1995). The fact that this result was not found by Lins et al. (2017) when studying US firms indicates that investors in the US reward companies for improving their CSR as much as possible. By contrast, European investors reward companies for avoiding the worst CSR scores but do not appreciate further CSR efforts unless it propels the company to the top. This finding could have an impact on companies' regional policies and level of CSR investment.

Table 3.3. Subprime Crisis-Period Returns and CSR Scores

This table presents regression estimates of Subprime crisis-period returns on CSR scores and control variables. Crisis-period returns are measured as both raw buy-and-hold returns and abnormal returns over the period August 2008 to March 2009. In Panel A, a linear measure of *CSR Score* is used, which is the rescaled average of Refinitiv's Environment and Social Pillar Scores measured at the end of 2006. In Panel B, we use dummy variables for *CSR Score*-quartiles such that *CSR Score Q2* takes the value of one if the firm is in the second *CSR Score*-quartile and zero otherwise, *CSR Score Q3* takes the value of one if the firm is in the third *CSR Score*-quartile and zero otherwise, and *CSR Score Q4* takes the value of one if the firm is in the fourth *CSR Score*-quartile and zero otherwise. For comparison, Columns (5) in Panels A and B show Lins et al.'s (2017) US Subprime Crisis results when explaining Crisis-Period Abnormal Returns with all controls added. Panel A, Column (5) is in Table II, Panel A, Column (4) of Lins et al. (2017); Panel B, Column (5) is in Table II, Panel A, Column (4) of Lins et al. (2017). The control variables are as defined in Table 3.1. Sector dummies are defined according to GICS. Lins et al.'s (2017) findings, quoted here in Columns (5) of Panels A and B, employ differently defined industry dummies instead of sector dummies. Country dummies are defined by country of headquarters. Financial firms and micro-cap stocks with a market capitalisation below EUR 250 million are removed from the sample. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Panel A: CSR Score: Raw and Abnormal Returns

Subprime Crisis: August 2008 – March 2009. N = 486

	(1) *Crisis Raw Return*	(2) *Crisis Abnormal Return*	(3) *Crisis Raw Return*	(4) *Crisis Abnormal Return*	(5) *LST '17 Abn. Return*
CSR Score	0.0583**	0.0752***	0.0507*	0.0504*	0.087***
	(0.0292)	(0.0285)	(0.0302)	(0.0301)	(0.032)
Ln(Market Cap)			0.00980	0.0184**	-0.015
			(0.00819)	(0.00830)	(0.011)
Long-Term Debt			-0.127	-0.105	-0.102
			(0.0798)	(0.0816)	(0.086)
Short-Term Debt			-0.399**	-0.451***	-0.384*
			(0.158)	(0.161)	(0.219)
Cash Holdings			0.0927	0.130	0.380***
			(0.118)	(0.119)	(0.091)
Profitability			0.327**	0.313**	0.732
			(0.128)	(0.130)	(0.509)
Book-to-Market			-0.0552***	-0.0467***	-0.045
			(0.0140)	(0.0141)	(0.058)
Negative B/M			0.0249	0.0329	0.049
			(0.101)	(0.0808)	(0.127)
Momentum			-0.177***	-0.122**	-0.285
			(0.0572)	(0.0584)	(0.044)
Idiosyncratic Risk			3.911	1.563	-8.870***
			(3.315)	(3.679)	(1.719)
Constant	-0.309***	-0.202***	-0.519**	-0.650***	0.0867***

	(0.0916)	(0.0659)	(0.239)	(0.232)	(0.032)
Four-factor loadings	Yes	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes	-
N	486	486	486	486	1,673
Adjusted R-squared	0.193	0.233	0.269	0.300	0.37

Panel B: Dummies for Quartiles of CSR Score: Raw and Abnormal Returns

Subprime Crisis: August 2008 – March 2009. *N* = 486

	(1) *Crisis Raw Return*	(2) *Crisis Abnormal Return*	(3) *Crisis Raw Return*	(4) *Crisis Abnormal Return*	(5) *LST '17 Abnormal Return*
CSR Score Q2	0.0582**	0.0600**	0.0456*	0.0413	0.0480
	(0.0279)	(0.0283)	(0.0266)	(0.0271)	(0.0341)
CSR Score Q3	0.0453	0.0501*	0.0288	0.0258	0.0562
	(0.0281)	(0.0289)	(0.0277)	(0.0285)	(0.0361)
CSR Score Q4	0.0655**	0.0789***	0.0580*	0.0539*	0.0727**
	(0.0289)	(0.0281)	(0.0298)	(0.0293)	(0.0362)
Ln(Market Cap)			0.0101	0.0190**	-0.012
			(0.00816)	(0.00835)	(0.011)
Long-Term Debt			-0.123	-0.101	-0.103
			(0.0802)	(0.0821)	(0.087)
Short-Term Debt			-0.379**	-0.433***	-0.386*
			(0.160)	(0.163)	(0.219)
Cash Holdings			0.0903	0.129	0.381***
			(0.117)	(0.119)	(0.091)
Profitability			0.327**	0.313**	0.734
			(0.128)	(0.130)	(0.508)
Book-to-Market			-0.0551***	-0.0465***	-0.046***
			(0.0142)	(0.0142)	(0.057)
Negative B/M			0.0140	0.0222	0.046
			(0.107)	(0.0852)	(0.058)
Momentum			-0.182***	-0.127**	-0.288
			(0.0571)	(0.0586)	(0.044)
Idiosyncratic Risk			4.088	1.680	-8.934***
			(3.302)	(3.658)	(1.729)
Four-factor loadings	Yes	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes	-
N	486	486	486	486	1,673
Adjusted R-squared	0.194	0.233	0.269	0.299	0.36

3.5.2 The Sovereign Debt Crisis: The reversed impact of having high CSR scores

We now turn to the analysis of the 2011 European Sovereign Debt Crisis. It is reasonable to suggest that a shock to trust in governments leads to reduced trust in other parts of society, including the financial sector (Leonardi et al., 2001). If European markets indeed experienced a secondary shock to trust in 2011, we would expect to find similar results as during the Subprime Crisis.

For our Sovereign Debt Crisis sample, however, we find a negative association between CSR scores and financial performance. The results are reported in Table 3.4. The coefficient on the continuous *CSR Score* variable, as seen in Panel A, is negative in all Sovereign Debt Crisis specifications, and significant at the 10% level in Column (3) and (4), with magnitudes of -0.0425 and -0.0416, respectively. One standard deviation in *CSR Score* (0.37) is associated with a raw and abnormal loss of 1.6% and 1.5%, respectively, on average, over the six-month period. The quartile analysis in Panel B confirms this negative association, as the coefficient on the best *CSR Score*-quartile, *CSR Score Q4*, is negative in all models and significantly so at the 10%-level in Columns (3) and (4).

The reason that the relationship between CSR scores and stock returns changes sign from the Subprime Crisis to the Sovereign Debt Crisis may be the result of the two crises being structurally different, originating in different ways, and impacting firm operations differently. However, this does not help to understand why CSR has a significantly negative effect on returns during the Sovereign Debt Crisis as opposed to being insignificant, for instance.

It could be argued that there was no shock to trust during the Sovereign Debt Crisis in the way there was during the Subprime Crisis. As previously outlined, CSR might only be expected to positively contribute to returns when systematic trust in financial markets is low. Perhaps the Sovereign Debt Crisis was perceived to merely be an ‘aftershock’ to the 2008 crisis. Moreover,

while the Subprime Crisis had fraudulent and negligent corporations at its core, the Sovereign Debt Crisis centred around struggling governments. That said, the impact of the Sovereign Debt Crisis on financial markets was indeed substantial, with an average stock return of -20.7% across companies in our sample during the second and third quarters of 2011. Even supposing that there was no significant crisis of trust during this period, the expectation would be for the value added by CSR to decrease, not for CSR to become harmful. We explore explanations for this in the next section.

[Table 3.4 starts on the next page.]

Table 3.4. Sovereign Debt Crisis-Period Returns and CSR Scores

This table presents regression estimates of Sovereign Debt crisis-period returns on CSR scores and control variables. Crisis-period returns are measured as both raw buy-and-hold returns and abnormal returns over the period April 2011 to September 2011. In Panel A, a linear measure of *CSR Score* is used, which is the rescaled average of Refinitiv's Environment and Social Pillar Scores measured at the end of 2009. In Panel B, we use dummy variables for *CSR Score*-quartiles such that *CSR Score Q2* takes the value of one if the firm is in the second *CSR Score*-quartile and zero otherwise, *CSR Score Q3* takes the value of one if the firm is in the third *CSR Score*-quartile and zero otherwise, and *CSR Score Q4* takes the value of one if the firm is in the fourth *CSR Score*-quartile and zero otherwise. The control variables are as defined in Table 3.1. Sector dummies are defined according to GICS. Country dummies are defined by country of headquarters. Financial firms and micro-cap stocks with a market capitalisation below EUR 250 million are removed from the sample. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Panel A: CSR Score: Raw and Abnormal Returns

Sovereign Debt Crisis: April 2011 – September 2011. N = 486

	(1) *Crisis Raw Return*	(2) *Crisis Abnormal Return*	(3) *Crisis Raw Return*	(4) *Crisis Abnormal Return*
CSR Score	-0.0233	-0.0222	-0.0425*	-0.0416*
	(0.0235)	(0.0233)	(0.0235)	(0.0233)
Ln(Market Cap)			0.0220***	0.0219***
			(0.00675)	(0.00674)
Long-Term Debt			-0.00290	-0.00168
			(0.0571)	(0.0568)
Short-Term Debt			0.0587	0.0658
			(0.143)	(0.142)
Cash Holdings			0.0520	0.0553
			(0.0862)	(0.0858)
Profitability			0.432***	0.430***
			(0.126)	(0.127)
Book-to-Market			-0.0215**	-0.0214**
			(0.00938)	(0.00928)
Negative B/M			0.0258	0.0255
			(0.0405)	(0.0406)
Momentum			0.0697*	0.0675*
			(0.0406)	(0.0402)
Idiosyncratic Risk			0.390	0.362
			(1.790)	(1.800)
Constant	0.0226	0.0259	-0.607***	-0.602***
	(0.0701)	(0.0684)	(0.187)	(0.186)

Four-factor loadings	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
N	486	486	486	486
Adjusted R-squared	0.356	0.197	0.441	0.302

Panel B: Dummies for Quartiles of CSR Score: Raw and Abnormal Returns

Sovereign Debt Crisis: April 2011 – September 2011. *N* = 486

	(1) *Crisis Raw Return*	(2) *Crisis Abnormal Return*	(3) *Crisis Raw Return*	(4) *Crisis Abnormal Return*
CSR Score Q2	-0.00601	-0.00534	-0.000243	0.000219
	(0.0225)	(0.0223)	(0.0215)	(0.0213)
CSR Score Q3	-0.0278	-0.0267	-0.0313	-0.0304
	(0.0240)	(0.0237)	(0.0226)	(0.0224)
CSR Score Q4	-0.0220	-0.0210	-0.0418*	-0.0410*
	(0.0239)	(0.0238)	(0.0243)	(0.0241)
Ln(Market Cap)			0.0227***	0.0226***
			(0.00677)	(0.00677)
Long-Term Debt			-0.00575	-0.00454
			(0.0568)	(0.0566)
Short-Term Debt			0.0654	0.0724
			(0.143)	(0.142)
Cash Holdings			0.0563	0.0597
			(0.0866)	(0.0862)
Profitability			0.417***	0.416***
			(0.128)	(0.129)
Book-to-Market			-0.0216**	-0.0214**
			(0.00924)	(0.00914)
Negative B/M			0.0168	0.0165
			(0.0429)	(0.0430)
Momentum			0.0733*	0.0710*
			(0.0410)	(0.0406)
Idiosyncratic Risk			0.459	0.430
			(1.742)	(1.752)
Four-factor loadings	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
N	486	486	486	486
Adj. R-squared	0.355	0.195	0.441	0.302

3.6 Controversy Avoidance as a Measure of CSR

In the context of CSR, 'window-dressing' is a term used to describe how a firm may pursue seemingly responsible initiatives to distract from its ethical failings (Connors et al., 2017). We put forth that a firm might try to improve its CSR score if it anticipates a future company-level challenge or crisis, such as litigation or scandal. We suggest that it is likely for this practice to have become more commonplace after the Subprime Crisis as ESG controversies became more pertinent to investors. Assuming investors became aware of this 'pre-emptive rehabilitation' trend, this would go a long way towards explaining why CSR scores were positively associated with returns during the Subprime Crisis but negatively so later on.

To investigate this, we use Refinitiv's ESG Controversies Scores which are calculated yearly on the basis of 23 ESG controversy topics taking into account a variety of sustainability and conduct issues which could lead to disputes and fines (Refinitiv, 2021). For example, one of the most prominent categories is Product Responsibility Controversies. This captures concerns regarding customers' health and safety as well as their privacy. Another featured category is Resource Use Controversies, which considers the company's environmental impact on natural resources and local communities. The penalty the company faces can be greater if the controversy leads to legal concerns. For instance, from a Workforce Controversies perspective, concerns regarding the unfair treatment of employees, including discrimination, harassment, and low compensation, are captured. Importantly, the controversy score would be even worse if employees opted to strike or initiate an industrial dispute. By looking at a wide variety of categories, these controversy scores strive to capture society's sentiments towards firms. In terms of construction, the scores are a media-based measure. The more media exposure a firm receives about its involvement in controversial dealings, the worse its score (Refinitiv, 2021).

There is little research on Refinitiv's controversy scores, and the findings are somewhat contradictory. For instance, De Franco (2020) shows that high-controversy stocks

underperform less controversial firms. However, Aouadi and Marsat (2018) find that ESG controversies are associated with higher firm value. For our testing, we construct a *Controversy Avoidance* variable by rescaling the Refinitiv ESG Controversies Score to range from -1 to +1.[14] The best score of +1 would indicate complete avoidance of controversy while the poorest score of -1 would indicate high involvement in controversial events. Using this scale has the benefit of being comparable to our *CSR Score* range with the same bounds used in Sections 3.3 and 3.4.

The merit of these scores lies in their objectiveness: while a firm can, to an extent, remedy its CSR score through investment in social initiatives, it is less able to window-dress its controversy score. The latter can only be improved through genuine, sustainable prudence, as well as avoidance of litigation through rigorous compliance, as measured by Refinitiv's 23 indicators. The media reporting on a company's ESG failings is either much less or not at all in the company's hands. One may argue that the CSR scores contain within them the level of firm controversy, making the latter redundant. However, we maintain that this is not the case for two main reasons. Firstly, the scores are offered separately by the data provider.[15] Again, this is probably because ESG scores are based on a variety of factors including many that the company can influence, such as the adoption of new policies. Conversely, controversy is based on external events beyond the company's direct influence, such as a fine being brought against it.

Secondly and more pertinently, we find that the two scores can be drastically different. In our Subprime Crisis sample, 41% of firms in the best *CSR Score* quartile belong to the worst

[14] As before, this is done by subtracting 50 from the percentile score, equating a score of 50/100 to 0, then dividing by 100.

[15] Refinitiv also offers an "ESG Combined Score" which imposes a penalty on the raw "ESG Score" if a company suffers from high controversy. We do not use this combined score because it does not capture the full variation in controversy: it is only lower than the raw scores for a firm if the firm's controversy is above a certain threshold. Otherwise, the combined score is equal to the raw score (Refinitiv, 2021).

Controversy Avoidance quartile. This proportion is even larger for the Sovereign Debt Crisis sample at 53%. We find that the correlation between a firm's *CSR Score* in one year and its *Controversy Avoidance* in the next has been negative and significant in every pair of years from 2005/6 until 2017/18, indicating that a high *CSR Score* somewhat predicts high controversy.[16] Notably, the correlation between 2006 *CSR Score* and 2007 *Controversy Avoidance* leading up to the Subprime Crisis is -25% while the correlation between 2009 *CSR Score* and 2010 *Controversy Avoidance* prior to the Sovereign Debt Crisis is a steeper -38%, significantly stronger at the 5% level. This means that, in 2009, a high *CSR Score* was a more reliable signal of high future controversy compared to three years prior.[17]

Figure 3.2 shows companies with a high $CSR\ Score_t$ (above 0.5) and low $Controversy\ Avoidance_{t+1}$ (below -0.5) for each crisis window, using a common sample of 474 firms.[18] We can clearly see that the number of companies with such a combination of CSR and controversy avoidance scores is much larger in the Sovereign Debt Crisis sample. Since the structure of these scores has not fundamentally changed between the crises, it is more likely that the intensifying correlation between *CSR Score* and year-ahead controversy is due to an underlying change in company behaviour in anticipation of troubled times. There is evidence that distressed firms increase investment in less costly ESG dimensions in the face of a hardship to protect their CSR image (Mecaj and Bravo, 2014). This is in line with our observation. Put differently, it seems that firms with knowledge of their impending distress try to pre-emptively rehabilitate their CSR scores. This is often referred to as the 'insurance mechanism' for CSR, whereby CSR is used to protect against future controversy (Kang et al., 2016). Under this

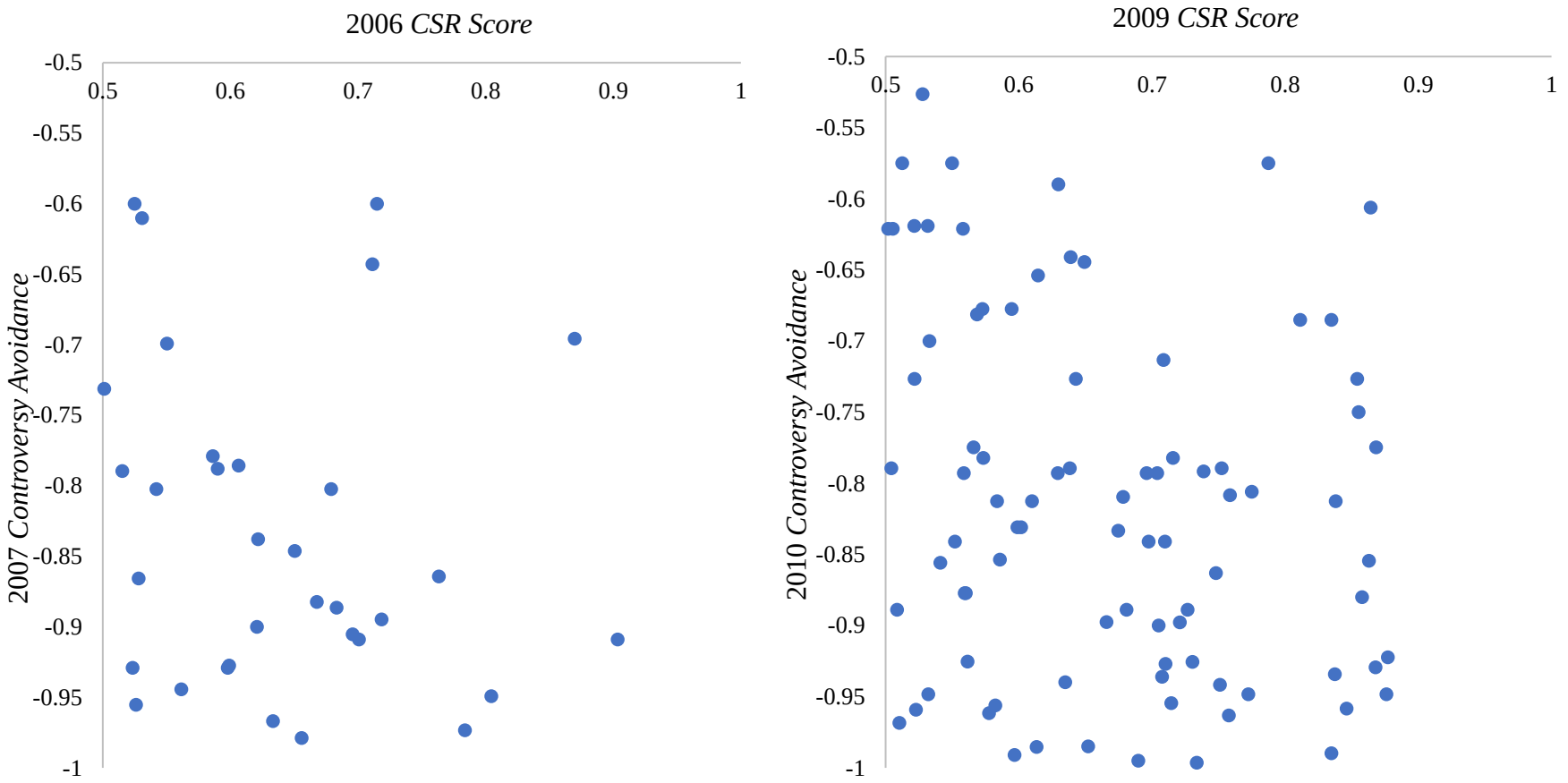
[16] We reiterate that high ESG controversy is captured by a low *Controversy Avoidance* score. As such, the less controversies a firm has, the higher its *Controversy Avoidance score*.
[17] Another reason we do not employ Refinitiv's combined score is because it measures ESG and controversy in the same year when we require year-ahead controversy.
[18] We lose 12 firm observations from our previous sample described in Table 3.1 for lack of available Refinitiv ESG Controversies Scores.

supposition, the companies in Figure 3.2 would be the ones exercising this practice most sharply. The density of the Sovereign Debt Crisis plot to the right is higher, meaning that more companies had high CSR scores despite high controversies in the later crisis, implying that this 'pre-emptive rehabilitation' was more commonplace in 2009 than in 2006.

[Figure 3.2. is on the next page.]

Figure 3.2. Companies with High *CSR Score*$_t$ and Low *Controversy Avoidance*$_{t+1}$

We include from the original sample of 486 firms only the 474 companies which have ESG Controversies data from the Refinitiv ESG database as of year-ends 2007 and 2010. *Controversy Avoidance* is the Refinitiv ESG Controversies score rescaled to a value between -1 and +1. *Controversy Avoidance* is measured at the end of 2007 and 2010 for the Subprime Crisis and Sovereign Debt Crisis, respectively. *CSR Score* is the 2006 and 2009 CSR scores for the Subprime Crisis and Sovereign Debt Crisis, respectively. This figure shows the companies with a high *CSR Score*$_t$ (above 0.5) and low *Controversy Avoidance*$_{t+1}$ (below -0.5) for each crisis window, side-by-side. The number of companies with such a combination of CSR and controversy avoidance scores is much larger in the Sovereign Debt Crisis window.



To better understand the properties of the *Controversy Avoidance* score, we consider its descriptive statistics in Table 3.5. We note that firms have attracted more controversy in 2010 than in 2007, with the mean *Controversy Avoidance* score falling considerably. This is almost identical in magnitude to the increase in the *CSR Score* over the same period of time, as reported in Table 3.1, Panel A. Looking at Table 3.5, Panel B, we see the negative correlation between *Controversy Avoidance* and the previous year's *CSR Score*. The table also reiterates that this is stronger leading up to the more recent crisis.

[Table 3.5. is on the next page.]

Table 3.5. Controversy Avoidance Descriptive Statistics

This table shows the descriptive statistics for a common subsample taken from the full sample of 486 firms used in Sections 3.3, 3.4, and 3.5 whose data is described in Table 3.1. We include from the original sample only the 474 companies which have ESG Controversies data from the Refinitiv ESG database as of year-ends 2007 and 2010. *Controversy Avoidance* is the Refinitiv ESG Controversies Score rescaled to a value between -1 and +1. *Controversy Avoidance* is measured at the end of 2007 and 2010 for the Subprime Crisis and Sovereign Debt Crisis, respectively. *CSR Score*$_{t-1}$ is the 2006 and 2009 CSR scores for the Subprime Crisis and Sovereign Debt Crisis, respectively. The measures in the first column of Panel B correspond to the respective measures for each crisis period. 2007 *Controversy Avoidance* is correlated with the Subprime Crisis measures, similarly to *CSR Score* in Table 3.1, Panel B (i); 2010 *Controversy Avoidance* is correlated with the Sovereign Debt Crisis measures, similarly to *CSR Score* in Table 3.1, Panel B (ii). All other definitions and data treatments are as described in Table 3.1.

Panel A: Summary Statistics, *N* = 474

	Mean	*SD (Std Dev)*	25th perc.	Median	75th perc.
2007 *Controversy Avoidance*	-0.039	0.434	0.018	0.174	0.207
2010 *Controversy Avoidance*	-0.113	0.463	-0.645	0.166	0.219

Panel B: Correlation with Other Variables from Respective Crisis Sample, *N* = 474

	2007 *Controversy Avoidance* (Subprime Crisis)	2010 *Controversy Avoidance* (Sovereign Debt Crisis)
CSR Score$_{t-1}$	-0.25	-0.38
Crisis Raw Return	-0.07	-0.02
Crisis Abnormal Return	-0.03	0.02
Ln (Mkt Cap)	-0.30	-0.45
Long-Term Debt	0.02	0.03
Short-Term Debt	0.01	-0.02
Cash Holdings	0.03	-0.02
Profitability	0.03	0.07
Book-to-Market	0.01	-0.08
Negative B/M	0.05	0.09
Momentum	-0.06	0.02
Idiosyncratic Risk	0.05	0.13

The fact that the 2009 *CSR Score* employed in our Sovereign Debt Crisis analysis (Table 3.4) is so strongly negatively correlated with 2010 *Controversy Avoidance* offers a potential explanation for why *CSR Score*'s association to performance is negative and significant in that sample. Perhaps investors began to consider the avoidance of ESG controversy more closely

in evaluating true corporate responsibility and making their investment decisions. There is some support for this in the increasing correlations between *Controversy Avoidance* and performance from the Subprime Crisis to the succeeding one. To test our hypothesis regarding investors' demand for low controversy, we repeat the regression analysis conducted in Section 3.5 with *Controversy Avoidance* in place of *CSR Score*.

The results are reported in Table 3.6.[19] The positive 10%-significant coefficient of *Controversy Avoidance* in Column (4) shows that firms with higher controversy avoidance significantly outperform firms with poor controversy avoidance. Looking at Panel B with the *Controversy Avoidance* quartiles, coefficients are positive for all dummies under all tests, indicating it is always better to avoid the highest-controversy quartile. However, the only significant case (10%-level) is with abnormal returns for the best quartile of companies which have little to no controversies, *Controversy Avoidance Q4*.

This is in line with the previously discussed literature which shows that controversy can negatively impact a firm's financial performance. A direct channel for this could be the reduction of profits on account of higher litigation costs (Murphy and Coombes, 2009). Meanwhile, an indirect channel is through the harm controversy does to a firm's reputation, leading to financial loss (Graham et al., 2008; Guiso et al., 2006).

[Table 3.6 starts on the next page.]

[19] Due to data shortage for the Refinitiv ESG Controversies score, we lose 12 observations from our previous analyses. Rerunning the baseline regressions reported in Tables 3.3 and 3.4 with this sample of 474 companies, instead of the full 486, does not change any of our previous findings.

Table 3.6. Sovereign Debt Crisis-Period Returns and ESG Controversy Avoidance

This table presents regression estimates of Sovereign Debt crisis-period returns on *Controversy Avoidance* and control variables. Crisis-period returns are measured as both raw buy-and-hold returns and abnormal returns over the period April 2011 to September 2011. In Panel A, a linear measure of *Controversy Avoidance* is used, which is the rescaled average of Refinitiv's ESG Controversies score measured at the end of 2010. In Panel B, we use dummy variables for *Controversy Avoidance*-quartiles such that *Controversy Avoidance Q2* takes the value of one if the firm is in the second *Controversy Avoidance*-quartile and zero otherwise, *Controversy Avoidance Q3* takes the value of one if the firm is in the third *Controversy Avoidance*-quartile and zero otherwise, and *Controversy Avoidance Q4* takes the value of one if the firm is in the fourth *Controversy Avoidance*-quartile and zero otherwise. The control variables are as defined in Table 3.1. Sector dummies are defined according to GICS. Country dummies are defined by country of headquarters. Financial firms and micro-cap stocks with a market capitalisation below EUR 250 million are removed from the sample. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

Panel A: Controversy Avoidance Score: Raw and Abnormal Returns

Sovereign Debt Crisis: April 2011 – September 2011. *N* = 474

	(1) *Crisis Raw Return*	(2) *Crisis Abn. Return*	(3) *Crisis Raw Return*	(4) *Crisis Abn. Return*
Controversy Avoidance	0.0191	0.0203	0.0305*	0.0324*
	(0.0153)	(0.0151)	(0.0180)	(0.0177)
Ln(Market Cap)			0.0201***	0.0205***
			(0.00763)	(0.00757)
Long-Term Debt			-0.00633	-0.00400
			(0.0570)	(0.0567)
Short-Term Debt			0.0557	0.0638
			(0.147)	(0.145)
Cash Holdings			0.0551	0.0566
			(0.0853)	(0.0844)
Profitability			0.412***	0.403***
			(0.130)	(0.129)
Book-to-Market			-0.0198**	-0.0192**
			(0.00998)	(0.00969)
Negative B/M			0.0429	0.0421
			(0.0419)	(0.0422)
Momentum			0.0770*	0.0747*
			(0.0419)	(0.0410)
Idiosyncratic Risk			1.924	1.860
			(1.818)	(1.811)
Constant	-0.0370	-0.0308	-0.608***	-0.613***
	(0.0545)	(0.0509)	(0.201)	(0.200)
Four-factor loadings	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
N	474	474	474	474
Adjusted R-squared	0.364	0.208	0.439	0.301

Panel B: Dummies for Quartiles of Controversy Avoidance: Raw and Abnormal Returns

Sovereign Debt Crisis: April 2011 – September 2011. *N* = 474

	(1) *Crisis Raw Return*	(2) *Crisis Abnormal Return*	(3) *Crisis Raw Return*	(4) *Crisis Abnormal Return*
Controversy Avoidance Q2	0.0175	0.0185	0.0232	0.0249
	(0.0199)	(0.0198)	(0.0208)	(0.0205)
Controversy Avoidance Q3	0.0230	0.0236	0.0337	0.0352
	(0.0202)	(0.0201)	(0.0218)	(0.0215)
Controversy Avoidance Q4	0.0225	0.0252	0.0387	0.0422*
	(0.0240)	(0.0235)	(0.0255)	(0.0248)
Ln(Market Cap)			0.0193**	0.0260***
			(0.00877)	(0.00855)
Long-Term Debt			-0.151*	-0.133
			(0.0803)	(0.0816)
Short-Term Debt			-0.408***	-0.463***
			(0.157)	(0.160)
Cash Holdings			0.0856	0.118
			(0.125)	(0.127)
Profitability			0.344**	0.327**
			(0.135)	(0.137)
Book-to-Market			-0.0460***	-0.0385***
			(0.0139)	(0.0142)
Negative B/M			0.0326	0.0342
			(0.0939)	(0.0775)
Momentum			-0.180***	-0.113*
			(0.0587)	(0.0594)
Idiosyncratic Risk			3.697	0.412
			(3.532)	(3.910)
Four-factor loadings	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
N	474	474	474	474
Adjusted R-squared	0.361	0.205	0.437	0.299

We run this test for the Subprime Crisis, too, using continuous and quartile *Controversy Avoidance* with and without controls for both raw and abnormal returns, and find all positive coefficients, but none are significant. This would imply that markets did not reward the avoidance of controversies as much back in 2008 as they did in 2011. As mentioned before,

this is intuitive because the Subprime Crisis itself could have urged investors to start worrying about controversies considering the serious market impact of corporate irresponsibility leading up to the 2008 crisis. In addition, public awareness about climate change and the role corporations play in environmental damage has been on the rise, so controversies surrounding global warming and pollution, measured by the Refinitiv ESG Controversies Scores, have become increasingly important.

It is also worth checking whether a firm's CSR score remains relevant in explaining performance after year-ahead controversy is accounted for. To test this, we create a *NetCSR* variable by regressing *CSR Score* on *Controversy Avoidance* and taking the residuals. We perform this separately for each crisis. This creates a CSR score measure that is statistically 'net' of controversies. We repeat the same baseline analysis done previously (Tables 3.3 and 3.4) replacing the raw *CSR Score* with *NetCSR*. In line with our findings, *NetCSR* loads almost identically to *CSR Score* under all models during the Subprime Crisis, with positive and significant coefficients. This indicates that the 2006 CSR scores were equally informative for investors before and after accounting for the element within them that predicts 2007 controversies. This would imply that the negative relationship between CSR scores and controversy avoidance was negligible to investors back in 2006. By contrast, *NetCSR* is entirely insignificant under all models for the Sovereign Debt Crisis. This supports that it was indeed the negative association between 2009 *CSR Score* and 2010 *Controversy Avoidance* that caused *CSR Score* to have a negative and significant coefficient in our previous analysis for the Sovereign Debt Crisis (Table 3.4).

3.7 Conclusion

Our results show that CSR impacts crisis-period returns differently in different geographical regions and during different times. There is already evidence that CSR becomes more relevant during times of financial crisis when trust in financial markets is low. This is due to CSR acting as a proxy for social capital, which gives rise to the scarce resource of societal trust. This creates the expectation that CSR would be a positive contributor to performance during any crisis of trust in the financial system.

Previous literature found this to be the case in the US during the 2008 Subprime Crisis. We test this hypothesis for two major European crises using a common sample of 486 European companies. We find that European markets also rewarded good CSR during the 2008 Subprime Crisis. However, while a quartile analysis for the US shows a monotonic relationship between CSR scores and performance, whereby belonging to a higher CSR quartile is always rewarded with better stock market returns, a similar analysis for our European Subprime sample reveals a different trend. While belonging to the best CSR quartile is associated with having the best returns, firms in the second-worst quartile saw better financial performance compared to those in the second-best quartile. While this may not seem intuitive, there is prior literature that supports non-linearity in the relationship between CSR and financial performance. For example, Barnett and Salomon (2012) find that having moderate CSR is more detrimental to financial performance than having low CSR, but having the best CSR is always most financially rewarding. From a consumer perspective, Meijer and Schuyt (2005) find that very low CSR triggers boycotts while pursuit of excellent CSR does not garner additional consumer interest. Our findings suggest that, at least during the Subprime Crisis, European market participants rewarded a firm's avoidance of the worst CSR quartile as well as its efforts in reaching the top CSR quartile but were less appreciative of moderately good CSR.

During the succeeding 2011 European Sovereign Debt Crisis, on the other hand, a negative association between CSR scores and crisis-period performance is found in our European sample. This is robust to alternative crisis-period start and end dates. Investigation into the CSR scores themselves reveals that they are negatively correlated with year-ahead ESG controversy avoidance, and more so around the Sovereign Debt Crisis. This means that firms with higher CSR scores in 2009 were more likely to be involved in ESG controversies in 2010. This could suggest that there was increased window-dressing of CSR scores following the Subprime Crisis. If this is driven by corporate decision making, our observation would fit the 'insurance mechanism' of CSR, whereby CSR is pursued to insure against forthcoming controversy (Kang et al., 2016).

However, if investors perceive inflated CSR scores as a precursor to controversy, a negative association between CSR scores and financial performance becomes plausible. Indeed, when we measure responsibility using controversy avoidance instead of raw CSR scores, the association between responsibility and returns becomes positive and significant during the Sovereign Debt Crisis. We argue that controversy avoidance is a more objective indicator of responsibility while raw CSR scores are vulnerable to firm manipulation. Li et al. (2019) discuss how CSR initiatives can be 'symbolic' versus 'substantive', but firms reporting on either can increase their CSR scores. Meanwhile, irresponsible behaviour is more likely to be reported by the media (Berkan et al., 2021; Dyck et al., 2010). Access to controversy-based measures of responsibility offers fertile ground for future research on the CSR-CFP link.

Another interesting avenue for future work would be to look at financial risk in place of return. The relationship between ESG and downside risk measures would be of particular interest during crisis windows when fear of loss is heightened. Existing research suggests that lower downside risk is associated with successful ESG engagement (e.g., Hoepner et al., 2018), higher ESG scores (e.g., Lööf and Stephan, 2019), and best-in-class ESG-screened portfolios

(e.g., Verheyden et al., 2016). These analyses could be expanded to explore whether this association is stronger during crisis windows.

Ultimately, with CSR scores on the rise, it appears firms heeded the lesson of the Subprime Crisis and became more mindful of their perceived responsibility in more recent years. However, it is also likely the case that investors, becoming more conscious towards ESG issues themselves, look more closely at several indicators of ESG excellence before making their socially responsible investment decisions. The eventual goal is to immunise measures of responsibility against manipulation as far as possible until they truly and unbiasedly capture a company's dedication to ethical behaviour and sustainable practices. Only with such a measure can we really determine the relationship between 'being good' and 'doing good'.

In the meantime, further research on how different markets and investors measure and reward CSR and how that changes over time is vital. This would not only improve our understanding of CSR and socially responsible investing as initiatives, but also impact the way companies and investors make decisions in different regions of the world and at different points in the business cycle.

4 WORK FLEXIBILITY AND FIRM PERFORMANCE DURING THE COVID-19 STOCK MARKET CRASH

4.1 Abstract

Using a sample of 2,532 North American and European firms, we test whether work flexibility is related to COVID-19 crisis-period performance. We find that firms whose employees were accustomed to work flexibility earned stock market returns 2 percentage points higher than their counterparts' during the first quarter of 2020. Grouping by business nature, we find that industries in which work flexibility is more commonplace significantly outperformed less work-flexible industries. We do not find evidence that work flexibility was value-enhancing prior to the pandemic. Our work flexibility measure captures whether companies provide their employees flexible working arrangements, including remote working, job sharing, and compressed work weeks. Since work flexibility is an initiative aimed to promote employee work-life balance, it is considered a workforce dimension of CSR. Our findings support the growing literature that CSR provides stock price resilience in the presence of an exogenous shock to returns. We isolate a single dimension of CSR that we posit is more contemporaneously relevant during the pandemic, thereby promoting the disaggregation of ESG measures.

4.2 Introduction

The COVID-19 pandemic has given rise to recessions globally, with stock markets around the world suffering great losses in the early months of 2020. A growing literature studies how COVID-19 stock returns could be explained by various financial and nonfinancial indicators (e.g., Ding et al., 2021), providing insight on what offers firms resilience during uncertain times. A stream of this research (e.g., Bae et al., 2021) explores the relationship between stock returns and CSR, which is thought to help a firm build social capital, earning it better performance during crisis periods (e.g., Lins et al., 2017). However, the findings of these studies are inconclusive, with the choice of sample, crisis window, CSR measurement, and control variables significantly altering results.

This heterogeneity legitimises concern over the aggregation of many different CSR indicators into multidimensional ESG scores. The use of aggregate ESG scores is thought to suppress possible links between specific CSR dimensions and equity market performance (Demers, 2021; Steffen, 2021). Depending on prevailing market conditions, some CSR dimensions can become more relevant than others.

Considering the widescale social distancing measures and travel restrictions given rise to by the COVID-19 pandemic, firms were forced to adapt their businesses to survive. Key stakeholders in the success of this effort were a firm's workforce. As such, in this chapter, we focus on a CSR dimension under the "Workforce" category of the Social Pillar of ESG[20]: work flexibility. We hypothesise that firms offering flexible work arrangements pre-2020 outperformed their counterparts during the COVID-19 recession. Using an international sample of over 2,500 firms in North America and Europe, we find support for our hypothesis.

[20] As per Refinitiv's categorisation of indicators in their ESG database framework.

The positive and significant relationship between work flexibility and stock market performance is robust to sector-level effects but not industry-level effects. We find that this is due to a strong correlation between performance and flexibility at the industry level, in line with previous work (e.g., Favilukis et al., 2020; Papanikolaou and Schmidt, 2020). These studies find that firms belonging to industries with higher work flexibility outperformed firms in less work-flexible industries. We complement the cited work by showing that this positive association is robust to the inclusion of controls for firm accounting and market-based variables and the use of an alternative work flexibility measure.

Furthermore, given that work flexibility has its advantages and disadvantages during calm periods (e.g., De Menezes and Kelliher, 2011), we suggest that it only becomes value-enhancing in the face of COVID-19 when it unexpectedly became vital. By conducting a difference-in-difference analysis, we find that work flexibility is significantly more value-enhancing during the first quarter of 2020, observed as the COVID-19 crisis, as compared to during the calm year prior.

We explore two further analyses. First, to address doubts that work flexibility could be a proxy for more general ESG performance, we consider whether a variety of aggregate ESG scores are better able to explain stock market performance, and we find this not to be the case. Finally, using interaction terms, we explore how different sectors valued work flexibility during our COVID-19 window. We find that, compared to all other sectors, work flexibility was most value-enhancing in Health Care, which is intuitive in light of the crisis being pandemic-driven.

Our study contributes to the literature in several ways. First, we contribute to the COVID-19 economic literature by explaining differences in stock market returns using an array of financial measures and a nonfinancial key variable of interest. We also contribute to the CSR literature by exploring the performance-responsibility link using the specific CSR dimension of work

flexibility. This adds to more recent work measuring CSR using isolated ESG dimensions (e.g., Garel and Petit-Romec, 2021) as opposed to aggregate, multidimensional ESG scores. Furthermore, in focusing on the responsibility-return link during a crisis period, we also contribute to the subset of literature relating CSR to social capital, believed to become more valuable during periods of high uncertainty.

Our contribution also extends to research in other disciplines. By considering work flexibility's value enhancement potential, we contribute to the limited literature on the benefits of flexible work arrangements, such as remote working, to the firm (e.g., Bloom, 2015; Sánchez et al., 2007). Our focus on a workforce-based ESG indicator also allows us to contribute to the growing literature relating supply-side frictions and financial market performance, particularly during recessions (e.g., Papanikolaou and Schmidt, 2020). Finally, since a firm's workforce forms part of its human capital (e.g., Shen et al., 2019), our study contributes to the research on the link between investment in intangible assets and crisis-period equity market performance (e.g., Demers et al., 2021).

The rest of this chapter is structured as follows. Section 4.2 summarises the motivation for this study by reviewing the relevant literature and describes the main hypotheses tested. Section 4.3 describes the sample and methodology employed in our analysis and reports initial findings of a positive correlation between work flexibility and performance. Section 4.4 presents our regression results where *Work Flexibility* is the key variable of interest, and Section 4.5 concludes.

4.3 Related Literature and Motivation of Study

4.3.1 Disaggregation of ESG Scores

In the CSR-CFP literature, aggregate ESG scores are often used to proxy for a firm's overall responsibility. More recently, however, numerous studies have isolated single CSR dimensions to test CSR hypotheses that are more 'particular' based on prevailing conditions and sentiment. During normal times, an average responsible investor may have no preference for one form of genuine CSR over another. However, as society's priorities change, investors may begin to seek the form of engagement that appears most contemporaneously relevant, hence most likely to provide a clear advantage to the firm in the present.

Following this intuition, we put forth that aggregating vastly different dimensions of CSR can potentially average away the contribution from the more relevant single dimensions, in line with Steffen (2021). This could explain why Bae et al. (2021) and others do not find a positive association between multidimensional ESG scores and returns during the pandemic-driven crisis. COVID-19 lockdowns placed a great emphasis on how corporations engage with their workforce and local communities, giving rise to a focus on the Social Pillar of ESG. For example, Shan and Tang (2020) find that Chinese firms whose employees were more satisfied had superior returns during COVID-19. We follow Shan and Tang (2020) in focusing on the social aspect since we believe it to be the most relevant during a pandemic-driven recession. This is because the social dimension of CSR focuses on a firm's commitment to maintaining the health and safety of the society in which it operates. Refinitiv (2021, p.25), for example, indicate that their Social Pillar Score reflects the company's effectiveness in providing employees with a "healthy and safe workplace" while maintaining equal opportunities. It also measures the company's commitment to "public health" and the integration of "customer health and safety" into operations (Refinitiv, 2021, p.25). These definitions help to clarify the link between a firm's social initiatives and how it is perceived during a public health crisis.

4.3.2 Work Flexibility

Also in line with Shan and Tang (2020), we look into the workforce dimension of the Social Pillar. More specifically, we consider the effect of work flexibility on the stock price performance of firms during the first quarter of 2020. We hypothesise that investors had greater confidence in the impending performance of firms whose employees were accustomed to flexible working. This is in expectation of an onslaught of travel and social restrictions which would see many workplace closures.

While figures vary based on samples and time of measurement, estimates suggest that a considerable portion of employees were working remotely throughout the COVID-19 pandemic. Remote working – also referred to as working from home (WFH), telecommuting or teleworking – increased drastically in light of social distancing measures. Bick et al. (2020) and Brynjolfsson et al. (2020) find that nearly half of the US workforce worked partly or completely from home by May 2020. Barrero et al. (2020) reach a similar conclusion and note that this represents a ten-fold increase from pre-pandemic levels. Beyond the US, figures from the Office for National Statistics indicate that 38% of the UK workforce was working remotely by mid-June 2020, months after the first lockdown (Elliott, 2020). An independent research platform estimates this to have been a larger 60% at the height of pandemic (Finder UK, 2020).

An emerging literature explores the feasibility of remote working across jobs, sectors, and regions, and well as its effect on various outcomes. Dingel and Neiman (2020) make use of the US database O*NET (Occupational Information Network) and determine that 37% of jobs in the US can be done remotely, with strong heterogeneity among cities and industries. Using a variant of Dingel and Neiman's (2020) WFH measure and focusing on worker-level outcomes, Mongey et al. (2020) find that employees in low-WFH jobs suffered greater job losses when COVID-19 restrictions began. This is in line with the findings of Lui and Mai (2020), Montenovo et al. (2020), and Kahn et al. (2020). Furthermore, Alon et al. (2020) find that

women were less protected from unemployment during the COVID-19 downturn on account of their lower participation in “telecommutable” (p. 7) sectors.

Our work is closest to the subset of studies considering how remote working has affected economic outcomes. For instance, Papanikolaou and Schmidt (2020) find that industries in which employees are more able to work remotely suffered milder losses during the COVID-19 crash. In line with Alon et al. (2020), the authors employ data from the American Time Use Survey and determine which occupations allow employees to work full days from home. Based on this, they construct an industry-level measure of COVID-19 work exposure which is higher the less the industry can telework. By regressing industry-level stock market returns on the continuous variable of work exposure, Papanikolaou and Schmidt (2020) find an economically and statistically significant negative association. Favilukis et al. (2020) also consider the relationship between telework and variation in asset prices during the COVID-19 recession. Using US data on job attributes, the authors determine an occupation’s telework flexibility to be higher when the occupation’s technological aspect is high and its need for person contact is low. These job-level scores are aggregated at the industry level, and industry quintiles based on telework flexibility are created. Favilukis et al. (2020) find that industries in the most-flexible quintile significantly outperformed those in the least-flexible quintile. The authors perform an out-of-sample test on the remaining G7 countries and find international support for their findings.

We contribute to this literature relating alternative work arrangements and equity market performance. Notably, we mediate the relationship by adding controls for firm market-based and accounting characteristics since there is abundant evidence that a variety of financial metrics – such as profitability, debt levels, and idiosyncratic risk – impact stock market performance (e.g., Bernanke and Gertler, 1989; Du et al., 2010; Goyal and Santa-Clara, 2003).

As such, to truly assess whether remote working was value-enhancing, the findings must be robust to the inclusion of controls, as implemented in our study.

Markedly, our work departs from the cited workplace flexibility research by using a measure of work flexibility that accounts for more than remote working. We define work flexibility in line with Refinitiv's "Working Hours Flexibility" ESG indicator employed in our study. Under this classification, a company provides work flexibility if it offers programmes or processes that help its employees achieve a good work-life balance. The arrangements considered by Refinitiv include not only remote working but also flexible working schedules, job sharing, and reduced or compressed work weeks.[21] We put forth that employees having been accustomed to any form of work flexibility, as measured by our indicator, helped them navigate pandemic-driven changes to their work life more seamlessly. Our indicator being retrievable from Refinitiv means that access to potentially restricted survey data is not necessary for determining which employees benefited from work flexibility. The availability of our indicator at the firm level also means that we do not need to assume uniform industry or job-level work flexibility.

Bal and Izak (2020) review and summarise into paradigms the many ways in which the academic literature has defined workplace flexibility. They find that it is not uncommon for studies to use the same terminology despite referring to different practices. Our Refinitiv-based characterisation of work flexibility corresponds to two of Bal and Izak's (2020) paradigms: flexible work and flexible work arrangements.[22] 'Flexible work' relates to the adaptation of employee contracts to allow greater flexibility in the face of changing circumstances. For

[21] As defined within the Thomson Reuters Eikon platform through which the Refinitiv ESG database was accessed for this thesis.

[22] Despite Bal and Izak's (2020) distinction between these practices, we use the terms "flexible work(ing)" and "flexible work arrangements" interchangeably with "work flexibility" in this thesis. This is because our measure from Refinitiv encompasses both practices and combinations thereof.

example, Wright and Bretthauer (2010) find that the employment of non-permanent agency nurses helps hospitals avoid undesirable overtime worked by permanent staff. The authors indicate that flexible work actually increases staffing costs, but the improved work environment helps with recruitment. This, in turn, helps to reduce the nurse shortage in the US healthcare system. Meanwhile 'flexible work arrangements' are practices relating to the timing and location of work, colloquially referred to as 'flextime' and 'flexplace', respectively (Hill et al., 2008). These practices gained popularity as a way of achieving work-life balance, so several studies consider whether employees find that to be true. The empirical research on the benefits of flexible work arrangements to employees reaches heterogenous findings (Allen et al., 2013). However, more recent industry-led research finds that flexible working is greatly appreciated by employees (e.g., CIPD, 2019).

The benefits and costs of work flexibility can also be considered from the firm's perspective. Benefits can come through tangible and intangible channels or an interplay of the two. Tangibly, firms allowing 'flextime' and 'flexplace' can save themselves money on overhead expenses by renting smaller offices and spending less on bills (Heathfield, 2021). Offering work flexibility has also been linked to higher employee satisfaction (e.g., US: McNall et al., 2009; Europe: Origo and Pagani, 2008) which in turn could help firms attract and retain the best talent. This means firms would have to spend less on advertising new vacancies and training new recruits. There is also evidence that employees who are allowed to work remotely are more productive than their in-office counterparts, appeasing concerns that lower employee monitoring would harm work quality (e.g., Bloom et al., 2015). Flexible working has also been associated with higher workforce diversity and improved employee engagement (CIPD, 2019), both of which have in turn been associated with improved employee performance (e.g., Elsaid, 2012, and Markos and Sridevi, 2010, respectively). Sánchez et al. (2007) find that teleworking firms are also more results-oriented which has positive effects on firm performance.

On the other hand, concerns have been raised as to whether there really is a business case for offering work flexibility, similarly to how CSR's value-enhancement is debated (De Menezes and Kelliher, 2011). This is because some of the cited benefits of work flexibility have been found to be contingent on the availability of other resources or programmes, such as continued supervisor support (Breaugh and Farabee, 2012). Telecommuting also increases cybersecurity risk, which can be costly for firms to mitigate (Anant et al., 2020). Indeed, our sample statistics summarised in Section 4.4 show that many firms still do not offer work flexibility as a by-default benefit. We further suggest that this is because certain industries, such as Manufacturing and Airlines, require more on-site employees due to the nature of their business.

With there being legitimate firm-level advantages and disadvantages to work flexibility, doubts arise as to whether it can enhance a firm's bottom-line during normal times. However, we suggest that, in the presence of an exogenous shock giving rise to high market uncertainty, the case for work flexibility being value-enhancing strengthens. This is in line with evidence that various other forms of flexibility, such as operating (Liu at al., 2021), financial (Fahlenbrach et al., 2020), and behavioural (Au et al., 2019), improve a firm's performance during crises periods, like the COVID-19 recession. We expect work flexibility to be no different, as it becomes more valuable in the face of work schedule disruptions given rise to globally by the pandemic. We contribute to the literature on the potential benefits of work flexibility to the firm by exploring whether its availability is value-enhancing before and during the COVID-19 crash. In doing so, we also contribute to the flexibility literature more broadly by considering whether work flexibility was able to protect stock prices in the presence of an exogenous shock.

4.3.3 Hypothesis Development

In early 2020, the worsening public health crisis meant that many firms had to quickly adapt to changing circumstances if they were to remain a going concern. It is reasonable to say that

firms which were prepared for flexible working, having had it on offer already, were better equipped to cope with expected widescale restrictions. This gives rise to our first hypothesis:

Hypothesis 1: Firms that offered their employees flexible working arrangements prior to the onset of the COVID-19 recession outperformed their counterparts during the crisis window.

We reason that investors expected milder business disruption for firms with pre-2020 work flexibility, ceteris paribus, and acted on those expectations early on. This would imply that the losses suffered by flexible firms were milder than their counterparts' just ahead of and during the peak of the COVID-19 crisis.

We also expect there to be a strong industry effect reflecting the fact that some industries are inherently more able to offer their employees flexible working and hence resume 'business-as-usual' during the pandemic. We expect investors to have reflected this in their decision making. As such, we hypothesise the following:

Hypothesis 2: Companies belonging to more flexible industries outperformed companies in less flexible industries during the COVID-19 recession.

In testing this relationship, we put forth that the reason some industries were able to perform better than others during the recession is in large part attributable to their ability and preparedness to work flexibly, in line with the findings of Papanikolaou and Schmidt (2020), and Favilukis et al. (2020). Our inter-industry analysis builds on these previous studies in three main ways. First, we account for flexible work arrangements beyond remote working. Second, we test the robustness of the relationship to the inclusion of controls for firm characteristics. Third, instead of using US-based definitions of occupation-level flexibility, we use an indicator available for individual firms and conduct the analysis on an international sample.

We also expect the market's reward for work flexibility to be higher during the COVID-19 recession as compared to during normal times. This gives rise to our third hypothesis:

Hypothesis 3: Work flexibility was more return-enhancing during the COVID-19 recession than in the calm year preceding.

Prior to pandemic-based restrictions becoming imminent, there was no clear, essential need for work flexibility as far as business operations were concerned, making the offer of work flexibility more two-sided. We hypothesise that the perceived downsides of flexible work arrangements were regarded more concernedly pre-COVID-19. These include fears over reduced employee monitoring, weakened cyber security, and a diminished teamwork culture (see e.g., De Menezes and Kelliher, 2011). During the pandemic, however, most companies would be forced to work flexibly to stay afloat, bringing closer the relationship between work flexibility, profitability, and stock performance. We employ a difference-in-difference panel to test this.

Finally, we suggest that work flexibility does not proxy for more general CSR engagement. We believe work flexibility to be a better indicator of COVID-19 crash performance compared to less tangible, more aggregate ESG measures, such as workforce welfare or multidimensional CSR. This is on account of aggregate ESG scores measuring too wide a diversity of dimensions (Demers et al., 2021; Steffen, 2021). As such, our fourth hypothesis is:

Hypothesis 4: Aggregate ESG measures do not explain COVID-19 crash performance superiorly to work flexibility.

We posit that the mechanism through which a firm's bottom-line can be protected is more immediately obvious with preparedness for flexible working. In exploring this, we contribute to the literature debating whether aggregate ESG scores explain differences in stock price performance during the COVID-19 crash. Our empirical testing is based on the sample and methodology described in the next section.

4.4 Data and Methodology

In determining whether a company offers flexible working, we rely on Refinitiv's "Working Hours Flexibility" dummy indicator, hereafter *Work Flexibility*. This shows whether a company offers working schedules which promote a work-life balance and considers arrangements such as remote work, job sharing, and compressed work weeks.[23] We construct our sample using companies for which Refinitiv has reported the flexibility indicator as of the end of 2018 within North America and Europe where we presume work flexibility, when offered, to be practiced wholesomely and effectively (e.g., Kulak and Tüzüner, 2020). We take work flexibility as of the end of 2019 where available, and 2018 otherwise (less than 20% of sample).[24] We are reassured in taking some 2018 data by the fact that less than 5% of firms for which both values are available show a change from one year to the next.[25]

Our core analysis endeavours to explain crisis-period holding return over the COVID-19 stock market crash using work flexibility and an array of control variables. We consider the crisis period to be the first quarter of 2020 when worldwide markets suffered the most, with high levels of synchronicity. While lockdowns and restrictions were introduced at different points in different countries, Figure 4.1 shows that developed equity markets reacted quite similarly in the early months of 2020. As such, to measure a firm's COVID-19 crisis-period

[23] As defined within the Thomson Reuters Eikon platform through which the Refinitiv ESG database was accessed for this thesis.

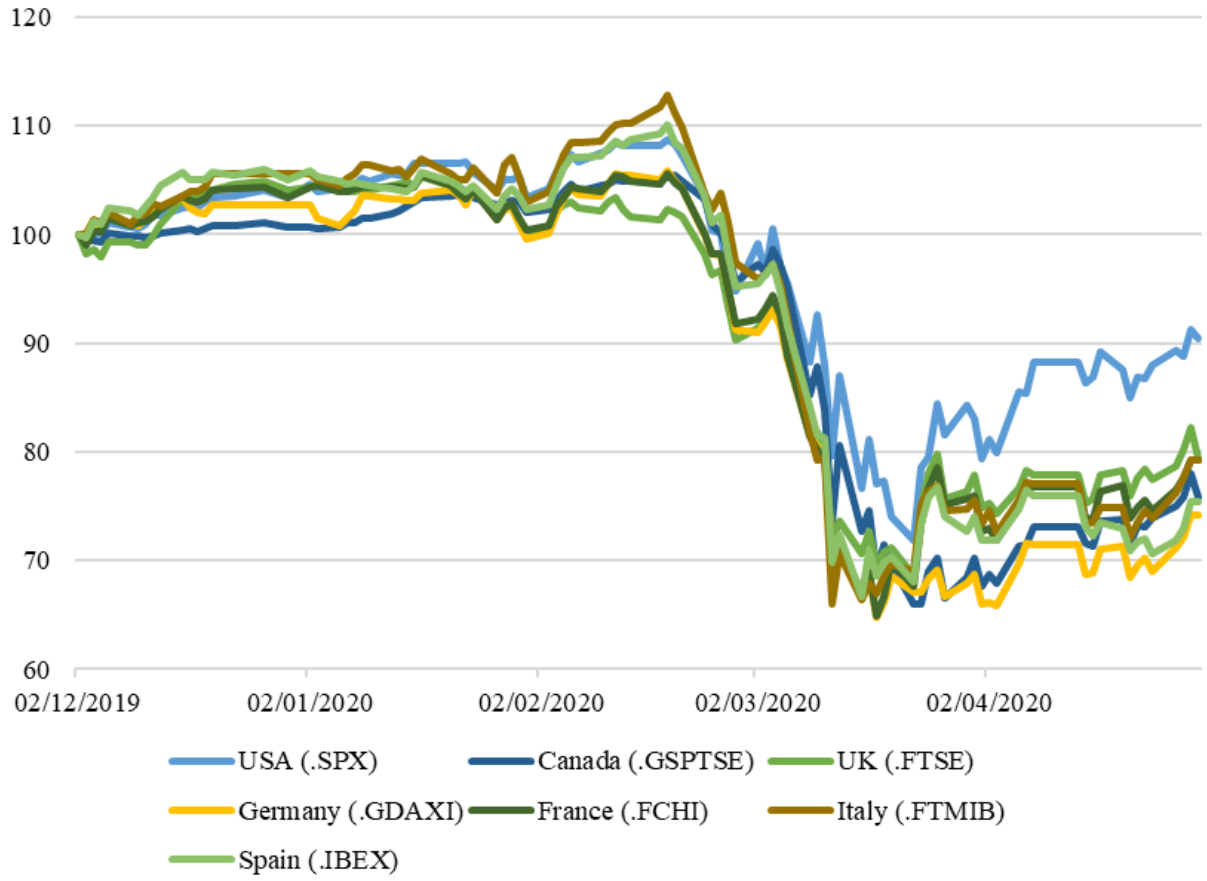
[24] For less than 20% of the firms in our sample, the 2019 indicator was not available at the time of the COVID-19 crisis. For these companies, the indicator as of 2018 was taken. There is good reason not to update this retrospectively, as our analysis implicitly assumes that investors would have been aware – at the start of 2020 – of whether a firm worked flexibly.

[25] In constructing our sample, we start with all firms for which the 2018 indicator is available. We require the 2018 indicator to maintain a common sample throughout the paper which considers not only a 2020 cross-sectional analysis but also a panel difference-in-difference analysis which extends backwards in time. Since the 2018 value is required for the panel, we select only the companies in possession of this value for *all* tests, allowing for more reliable comparison and deduction across analyses.

performance, we calculate its buy-and-hold stock return over the first quarter of 2020 in line with Demers et al. (2021).

Figure 4.1. Developed Markets' Equity Indices, 2 December 2019 = 100

The figure shows the value of a nominal 100 invested in each of the main equity market indices in the US (S&P500), Canada (S&P/TSX Composite Index), the UK (FTSE100), Germany (DAX), France (CAC), Italy (FTMIB) and Spain (IBEX) over the period December 2019 to April 2020. This figure shows the high synchronicity between the main equity indices of North American and European developed markets during the first quarter of 2020, adopted in our study as the COVID-19 crisis window. Price data was obtained from Refinitiv via Thomson Reuters Eikon.



The control variables encompass both accounting and market measures. The accounting variables are measured at the end of 2019, just prior to the COVID-19 crisis observation period. The firm characteristics captured are *Long-Term Debt*, *Short-Term Debt*, *Cash Holdings* and *Profitability*, all of which are included as a ratio to Total Assets. Furthermore, equity *Book-to-*

Market and a dummy for *Negative B/M* firms are included.[26] We also consider market-based indicators of risk and return. *Momentum* is measured as the firm's raw buy-and-hold return over the 12 months preceding the crisis period and is included to control for the firm's performance leading up to the crash. A firm's stock price volatility is controlled for with *Idiosyncratic Risk*, measured as the residual variance from market model regressions. Carhart (1997) four-factor loadings, based on regressions of monthly data for five years preceding the crisis, are also added to control for further effects. All non-dummy variables are winsorised at the 1st and 99th percentiles to mitigate the effect of outliers. All financial and accounting data is obtained from Refinitiv via Thomson Reuters Eikon (unless otherwise stated). We exclude companies for which the required accounting data is unavailable and/or those with fewer than 12 months of observations for market-model and four-factor regressions. This yields a final sample of 2,532 companies.

The descriptive statistics for this sample are summarised in Table 4.1. In Panel A, we note that fewer firms offer work flexibility compared to the contrary, but flexible firms still constitute a large portion of the sample at just over 37%. The *Crisis-Period Return* over the first quarter of 2020 is negative at all reported percentiles and has a mean of -32%. This is economically significant and indicative of investor concern over the performance of firms in the worsening pandemic. In Panel B, we see a positive correlation between crisis-period performance and work flexibility, which begins to provide support for our first two hypotheses. We also observe a very strong positive relationship between flexibility and size, as well as size and performance. This emphasises the need to control for size in our investigation of the flexibility-performance link.

[26] The latter corrects for the behaviour exhibited by negative book-to-market firms which are likely to be in distress and hence act more like high book-to-market firms (Fama and French, 1992).

Furthermore, we note that the correlation between *Work Flexibility* and *Momentum* is negative. This would imply that the performance of flexible firms was ever so slightly worse than their counterparts' over the year 2019 when the pandemic was not in sight. This lends support to our third hypothesis that work flexibility was not rewarded by investors prior to flexibility becoming more of a necessity in the face of COVID-19 restrictions. We consider this hypothesis in a panel setting in Section 4.5.3.

[Table 4.1 starts on the next page.]

Table 4.1. Descriptive Statistics: COVID-19 and Work Flexibility

The sample consists of 2,532 firms headquartered in North America and Europe with "Working Hours Flexibility" data available from the Refinitiv ESG database as of year-end 2018 and returns available during the period January to March 2020. *Work Flexibility* is a dummy with a value of one for firms that offer flexible work arrangements and zero otherwise, measured as of year-end 2019 where available and 2018 otherwise. *Crisis-Period Return* is the raw buy-and-hold return computed over the period January to March 2020. Accounting data is based on the last quarter of 2019 and is obtained from Refinitiv via Thomson Reuters Eikon. *Market Capitalisation* is in millions of USD. *Long-Term Debt* is computed as long-term debt divided by total assets. *Short-Term Debt* is computed as debt in current liabilities divided by total assets. *Cash Holdings* is computed as cash and marketable securities divided by total assets. *Profitability* is computed as operating income divided by total assets. *Book-to-Market* is computed as book value of equity divided by market value of equity. *Negative B/M* is a dummy variable set to one when the book-to-market ratio is negative and zero otherwise. *Momentum* is the raw return over the period January 2019 to December 2019. *Idiosyncratic Risk* is computed as the residual variance from the market model estimated over the five-year period ending in December 2019 using monthly data and the K. French Developed Markets value-weighted index as the market proxy. All financial data is obtained from Refinitiv via Thomson Reuters Eikon (unless otherwise stated). Control variables and returns are winsorised at the 1st and 99th percentiles.

Panel A: Summary Statistics, *N* = 2,532					
	Mean	*SD (Std Dev)*	25th perc.	Median	75th perc.
Work Flexibility	0.373	0.484	0.000	0.000	1.000
Crisis-Period Return	-0.321	0.218	-0.460	-0.313	-0.170
Market Capitalisation (US$m)	13989	31228	1262	3603	11112
Long-Term Debt	0.258	0.185	0.120	0.243	0.368
Short-Term Debt	0.034	0.053	0.001	0.013	0.047
Cash Holdings	0.103	0.117	0.025	0.064	0.136
Profitability	0.068	0.092	0.032	0.064	0.108
Book-to-Market	0.619	0.847	0.213	0.405	0.734
Negative B/M	0.038	0.192	0.000	0.000	0.000
Momentum	0.225	0.353	0.022	0.220	0.409
Idiosyncratic Risk	0.012	0.016	0.004	0.007	0.012

(Continued…)

Panel B: Correlation Matrix, *N* = 2,532										
	Work Flex.	*C.-P. Return*	*Ln (Mkt Cap)*	*LT Debt*	*ST Debt*	*Cash Hold.*	*Prof.*	*B/M*	*Neg. B/M*	*Mom.*
Crisis-Period Return	0.13									
Ln(Mkt Cap)	0.26	0.33								
Long-Term Debt	-0.05	-0.14	0.03							
Short-Term Debt	0.14	0.07	0.04	-0.05						
Cash Holdings	0.01	0.13	-0.01	-0.34	-0.06					
Profitability	0.04	0.15	0.32	-0.04	-0.08	0.06				
Book-to-Market	-0.03	-0.30	-0.42	-0.03	0.01	-0.13	-0.32			
Neg. B/M	0.03	-0.05	-0.05	0.35	0.09	0.00	0.05	-0.20		
Momentum	-0.02	0.25	0.28	-0.03	-0.06	0.08	0.13	-0.42	-0.04	
Idiosyncratic Risk	-0.11	-0.31	-0.50	0.11	0.02	0.10	-0.35	0.40	0.22	-0.23

Another imperative consideration in our analysis is business nature, captured by sector or industry categorisation. We expect there to be a strong relationship between business nature and flexibility, with some sectors being inherently more flexible than others. Throughout our study, we define sectors and industries as per the Global Industry Classification Standard (GICS). In our sample of 2,532 companies, we capture 11 GICS sectors: Communication Services, Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Information Technology, Materials, Real Estate, and Utilities. Across these 11 sectors, we can further classify our companies into 68 GICS industries.

From our sample, we can make anecdotal observations of industry ability to offer work flexibility and COVID-19 performance. For example, some industries, like Mining (Materials sector), are largely hindered from offering work flexibility while others, such as Telecommunication Services (Communication Services sector), are characteristically more flexible. Incidentally, we observe that Mining companies performed nearly three times as poorly on average as Telecommunication Services companies during the crisis. It is also understandable that some industries, like Household Products (Consumer Staples sector), weathered the storm of the crisis relatively well while others, like the Airlines industry (Industrials sector), were impacted more harshly; again, we note that the former of these is a significantly more flexible industry.

In light of this trend, we hypothesise that the positive 0.13 correlation between *Crisis-Period Return* and *Work Flexibility* reported in Table 4.1 is driven by higher-flexibility industries being ideally positioned to survive a pandemic-driven recession where movement and face-to-face interaction are restricted. To consider this, we first define an industry's level of work flexibility as the within-sample percentage of its companies that offer flexible working arrangements. We

also create within-sample equally weighted industry indices to approximate crisis-period performance at the industry-level.[27]

Figure 4.2, A. shows a scatter plot of these variables, where each point represents an industry and is sized in proportion to the industry's number of firms (i.e., a 'bubble' chart). The plot and estimated trendline indeed show that more flexible industries suffered milder losses in our observation window. In line with this representation, we find a positive and significant Pearson correlation between an industry's flexibility and the average performance of its companies (ρ=0.32, 1%-significant). We also consider this relationship using the broader classification of 11 sectors. Figure 4.2, B. shows the corresponding bubble chart for the sector-level variables, for which the Pearson correlation is also positive and significant (ρ=0.67, 5%-significant). Section 4.5 outlines the regression tests exploring the flexibility-performance relationship.

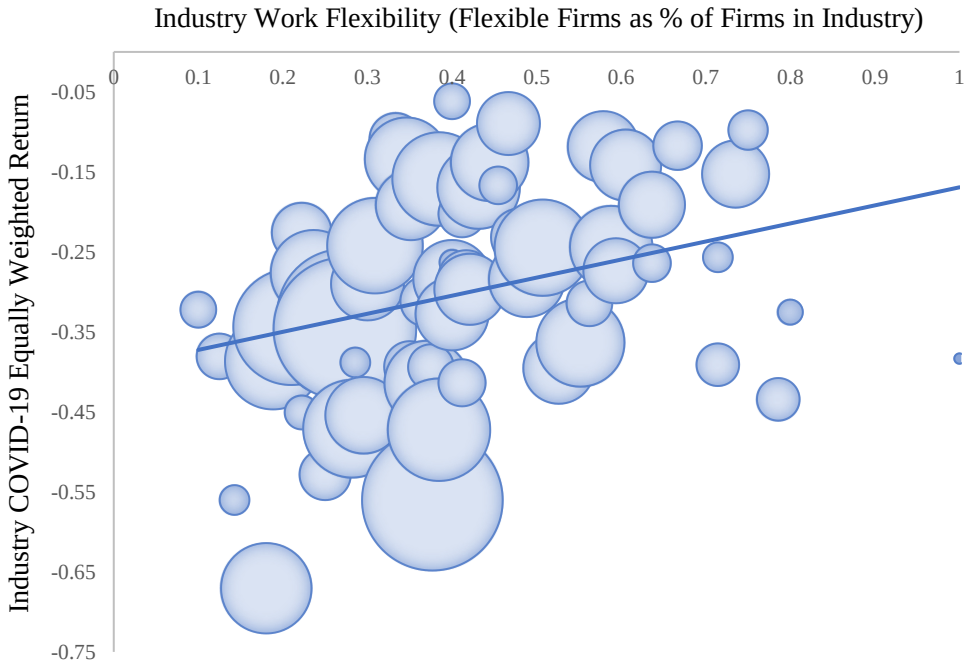
[Figure 4.2. is on the next page.]

[27] Our results are not materially different when using value-weighted indices and/or a value-weighted sector-level flexibility measure.

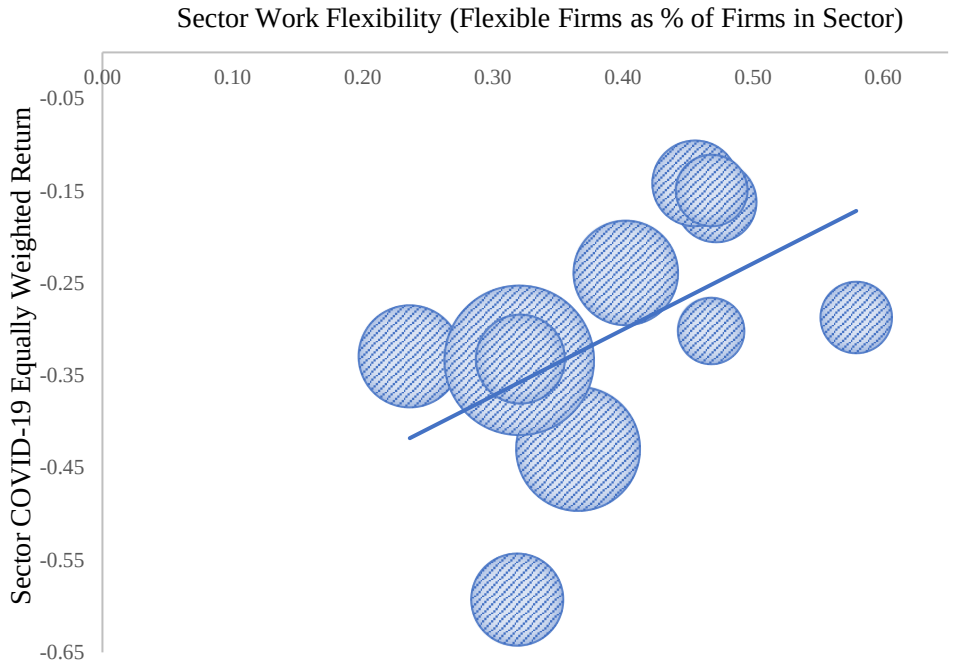
Figure 4.2. Industry/Sector Flexibility and COVID-19 Stock Performance

Figure 4.2 A. (B.) plots Industry (Sector) Work Flexibility against COVID-19 stock market performance. On the horizontal axis, Industry (Sector) Work Flexibility is the percentage of firms within an industry (a sector) which offered employees flexible working pre-2020, as measured by the "Working Hours Flexibility" indicator from the Refinitiv ESG database. On the vertical axis, Return is the equally weighted buy-and-hold return for the firms within the industry (A.) or sector (B.) computed over the period January to March 2020, observed as the COVID-19 crisis. Levels of flexibility and return are determined based on the firms in our sample, *N*=2,532. The data points in these 'bubble charts' are sized in proportion to the number of firms in the industry (A.) or sector (B.). Sectors and industries are determined according to the Global Industry Classification Standard (GICS).

A. Industry Work Flexibility and COVID-19 Performance



B. Sector Work Flexibility and COVID-19 Performance



4.5 Crisis-Period Returns and Work Flexibility

4.5.1 Baseline Analysis: Reward for flexibility at the firm level

We start examining the work flexibility and COVID-19 value-enhancement question by running a cross-sectional regression. The dependent variable, as described in the previous section, is the buy-and-hold return over the first quarter of 2020, which we consider the COVID-19 recession window in line with Demers et al (2021). We attempt to explain the differences in returns across our sample of 2,532 companies using the *Work Flexibility* dummy and a number of control variables, including those for country and sector or industry effects. For this test, the regression equation is as follows:

$$\textit{Crisis-Period Return}_i = \alpha + \beta_1 \textit{Work Flexibility}_i + \beta_2 \textit{Ln(Market Cap)}_i + \beta_3 \textit{Long-Term Debt}_i + \beta_4 \textit{Short-Term Debt}_i + \beta_5 \textit{Cash Holdings}_i + \beta_6 \textit{Profitability}_i + \beta_7 \textit{Book-to-Market}_i + \beta_8 \textit{Negative B/M}_i + \beta_9 \textit{Momentum}_i + \beta_{10} \textit{Idiosyncratic Risk}_i + \textit{Four-Factor Loadings} + \textit{Country Dummies} + \textit{Sector} \text{ or } \textit{Industry Dummies} + e_i \qquad \textit{(Equation 4.1)}$$

As previously discussed, there is a relationship between a firm's business nature and its likelihood of offering work flexibility to its employees. In this part of the analysis, we try to ascertain whether there is a market reward for work flexibility beyond sector or industry effects; therefore, we control for business nature in a variety of ways. First, we use sector dummies. The advantage of adding sector dummies instead of industry dummies is rooted in the number of additional regressors required. While there are 68 industries, there are only 11 sectors, so controlling for only sector-level effects means fewer degrees of freedom are lost. Also, there is less chance of unpredictable interaction between a large number of control dummies and the dummy variable of interest, *Work Flexibility*. However, for robustness and completeness, we also perform the test again with the full set of 68 industry dummies.

Second, we use a model demeaned at the sector level. We demean all non-dummy variables using only the companies in our sample to set sector averages. Using sector-adjusted variables allows us to observe return sensitivities that remain after sector effects are accounted for. Under this approach, we naturally do not add any business nature (sector or industry) dummies. We summarise the results of both methodologies in Table 4.2, where Columns (4) and (5) report the coefficients of the demeaned variables.

Table 4.2. COVID-19 Crisis-Period Returns and Work Flexibility

This table presents regression estimates of COVID-19 crisis-period returns on firm work flexibility and control variables. In Columns (1) – (3), *Raw Crisis-Period Returns* are measured as unadjusted buy-and-hold returns over the period January to March 2020. In Columns (4) and (5), all non-dummy variables, including returns on the left-hand side, are demeaned at the sector level, using equally weighted industry averages. We use a dummy variable for firm work flexibility obtained from Refinitiv's ESG database. Sectors and industries are determined according to GICS. The control variables are as defined in Table 4.1. Country dummies are defined by country of headquarters. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

				Demeaned Model	**Demeaned Model**
	(1)	(2)	(3)	(4)	(5)
	Raw Crisis-Period Return	*Raw Crisis-Period Return*	*Raw Crisis-Period Return*	*Sector-Adjusted Crisis-Period Return*	*Sector-Adjusted Crisis-Period Return*
Work Flexibility	0.0204**	0.0143*	0.00688	0.0329***	0.0142*
	(0.00902)	(0.00813)	(0.00805)	(0.00819)	(0.00795)
Ln(Market Cap)	0.0249***	0.0188***	0.0179***		0.0187***
	(0.00304)	(0.00290)	(0.00296)		(0.00289)
Long-Term Debt	-0.0998***	-0.140***	-0.134***		-0.140***
	(0.0274)	(0.0250)	(0.0255)		(0.0248)
Short-Term Debt	0.165*	0.0751	0.102		0.0767
	(0.0851)	(0.0759)	(0.0761)		(0.0757)
Cash Holdings	0.164***	0.122***	0.109***		0.122***
	(0.0395)	(0.0379)	(0.0400)		(0.0378)
Profitability	-0.0745	0.101**	0.122**		0.101**
	(0.0533)	(0.0514)	(0.0527)		(0.0514)
Book-to-Market	-0.0306***	-0.0118*	-0.00923		-0.0117*
	(0.00804)	(0.00700)	(0.00711)		(0.00698)
Negative B/M	0.00573	0.0253	0.0300		0.0254

	(0.0275)	(0.0256)	(0.0253)		(0.0256)
Momentum	0.0670***	0.0712***	0.0753***		0.0712***
	(0.0140)	(0.0135)	(0.0138)		(0.0135)
Idiosyncratic Risk	-1.887***	-0.886**	-0.918**		-0.893**
	(0.430)	(0.413)	(0.427)		(0.411)
Constant	-0.757***	-0.749***	-0.792***	-0.0309	-0.0182
	(0.0805)	(0.0834)	(0.0924)	(0.0588)	(0.0485)
Four-factor loadings	Yes	Yes	Yes	Yes	Yes
Sector/Industry Dummies	None	Sector	Industry	None	None
Country dummies	Yes	Yes	Yes	Yes	Yes
N	2,532	2,532	2,532	2,532	2,532
Adjusted R-squared	0.226	0.400	0.437	0.038	0.155

Under all models, the *Work Flexibility* dummy loads positively on crisis-period returns. While the magnitude of the coefficient varies between specifications, overall, it appears as though work flexibility was a positive return contributor during the first COVID-19 recession. This lends support to our first hypothesis suggesting that investors believed companies whose employees were accustomed to working flexibly would be more successful in the uncertain times to come.

Under the baseline specification in Column (1), where no sector or industry controls are added, the *Work Flexibility* coefficient suggests that, on average during the pandemic window, work-flexible firms earned stock market returns 2.04 percentage points greater than their non-flexible counterparts. However, accounting for sector or industry effects weans the importance of our *Work Flexibility* indicator, rendering it insignificant in Column (3). This is the model where the full 68 industry dummies are employed. This lack of significance emphasises the need to further explore the relationship between business nature and work flexibility. Investors could also make investment decisions based on industry analysis, foreseeing that more flexible industries are less likely to be disrupted in the face of incoming restrictions. This would mean that some of the return differential captured by

the industry dummies unfairly diminishes the estimate of *Work Flexibility*'s value enhancement.[28] To explore whether there is a relationship between industry performance and work flexibility, we amend the above analysis as described in the next section.

4.5.2 Grouping by Industry: Is business nature the key player?

We proceed to consider whether merely belonging to a work-flexible industry enhances a firm's COVID-19 return, in line with Papanikolaou and Schmidt (2020) and Favilukis et al. (2020). One way to test whether higher-flexibility industries outperformed lower-flexibility ones is to regress an industry's return on its level of work flexibility and a vector of control variables. However, this would confine us to a maximum of 68 observations, one for each GICS industry in our sample. To strengthen the power of our test, we maintain the use of our firm-level sample, containing 2,532 observations, and employ quartile dummies for industry-level work flexibility instead.

We start by placing our sample's 68 industries into four levels of work flexibility, creating four quartiles of roughly 17 industries each. As before, we define an industry's level of flexibility by the proportion of its firms, within our sample, that offer work flexibility. Companies belonging to any of the 17 most flexible industries, i.e., the top quartile of industries ranked by work flexibility, will have an *Industry Work Flexibility Q4* dummy value of one. The *Q4* dummy will be zero for all other companies. The *Industry Work Flexibility Q3* dummy will be one for companies belonging to the second-most flexible 17 industries and zero otherwise. The *Industry Work Flexibility Q2* dummy will be one for companies in the 17 second-least flexible industries and zero otherwise. Finally, the *Industry*

[28] The same logic can explain why the *Work Flexibility* coefficient is also weaned by the inclusion of sector dummies, but *Work Flexibility* remains significant when only the 11 sector dummies are employed, as per Column (2).

Work Flexibility Q1 dummy takes the value of one if a company is in one of the 17 least flexible industries and is zero for all other firms.

Naturally, the industries are not all of equal size, i.e., there is a different number of firms in each industry and hence in each quartile. We find that there are 389, 599, 640, and 904 firms in *Q4*, *Q3*, *Q2*, and *Q1*, respectively. While the middle quartiles, *Q3* and *Q2*, each represent roughly 25% of the firms in the sample, we note that fewer firms belong to highly flexible industries in *Q4* (15% of sample) as compared to highly inflexible ones in *Q1* (35% of sample).

In Section 4.5.1, we employed a single firm-level dummy for work flexibility, identifying whether the firm itself offers its employees flexible working or not. In this analysis, we replace this single dummy with the *Industry Work Flexibility Q2-Q4* dummies as follows:

$$\textit{Crisis-Period Return}_i = \beta_1 \textit{Industry Work Flexibility Q2} + \beta_2 \textit{Industry Work Flexibility Q3} + \beta_3 \textit{Industry Work Flexibility Q4} + \beta_4 \textit{Ln(Market Cap)}_i + \beta_5 \textit{Long-Term Debt}_i + \beta_6 \textit{Short-Term Debt}_i + \beta_7 \textit{Cash Holdings}_i + \beta_8 \textit{Profitability}_i + \beta_9 \textit{Book-to-Market}_i + \beta_{10} \textit{Negative B/M}_i + \beta_{11} \textit{Momentum}_i + \beta_{12} \textit{Idiosyncratic Risk}_i + \textit{Four-Factor Loadings} + \textit{Country Dummies} + e_i \quad \textit{(Equation 4.2)}$$

These quartile dummies no longer identify whether a firm singularly offers work flexibility, but rather whether the industry the firm belongs to generally does. This allows us to compare the performance of industry groups, clustered by their level of work flexibility, to each other. Constructing the test in this way permits a direct comparison between the performance of firms in one industry work-flexibility quartile and a subsequent one. The results of this test are reported in Table 4.3.

Table 4.3. COVID-19 Crisis-Period Returns and Industry Work Flexibility

This table presents regression estimates of COVID-19 crisis-period returns on industry-level work flexibility quartiles and control variables. *Crisis-Period Return* is measured as the raw holding-period returns over the period January to March 2020. We use dummy variables for industry-level work flexibility quartiles such that *Industry Work Flexibility Q2* takes the value of one if the firm is in the second-least-flexible 17 industries and zero otherwise, *Industry Work Flexibility Q3* takes the value of one if the firm is in the second-most flexible 17 industries and zero otherwise, and *Industry Work Flexibility Q4* takes the value of one if the firm is in the most flexible 17 industries and zero otherwise. Industries are determined according to GICS. The control variables are as defined in Table 4.1. Country dummies are defined by country of headquarters. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Crisis-Period Return*	(2) *Crisis-Period Return*
Industry Work Flexibility Q2	0.0339***	0.0253**
	(0.0113)	(0.0105)
Industry Work Flexibility Q3	0.0967***	0.0603***
	(0.0107)	(0.0103)
Industry Work Flexibility Q4	0.136***	0.104***
	(0.0124)	(0.0121)
Ln(Market Cap)		0.0219***
		(0.00297)
Long-Term Debt		-0.0998***
		(0.0271)
Short-Term Debt		0.145*
		(0.0840)
Cash Holdings		0.129***
		(0.0408)
Profitability		-0.0479
		(0.0535)
Book-to-Market		-0.0265***
		(0.00781)
Negative B/M		-0.00391
		(0.0270)
Momentum		0.0735***
		(0.0138)
Idiosyncratic Risk		-1.892***
		(0.424)
Four-factor loadings	Yes	Yes
Country dummies	Yes	Yes
N	2,532	2,532
Adjusted R-squared	0.720	0.763

The coefficients on *Industry Work Flexibility Q2*, *Q3*, and *Q4* allow us to see, as compared to belonging to one of the least work-flexible 17 industries[29], the incremental effect of a company being in the second-least flexible, second-most flexible, and most flexible 17 industries, respectively. The industry work flexibility quartile coefficients are always positive and increase monotonically, whether controls are excluded or included. All quartile coefficients are also highly significant. This indicates that, all else equal and accounting for a variety of other factors, firms belonging to more work-flexible industries performed better during the COVID-19 crisis period. This can be viewed as a 'free-rider' effect: on average, a firm belonging to a work-flexible industry benefited during the COVID-19 crash, irrespective of whether the firm itself offered work flexibility or not. This is not to say that there was no incremental effect for the firm being work-flexible within its industry.

We find that the same conclusion is reached if firm grouping is done at the sector level. As before, sector work flexibility is defined as the percentage of firms in the sector offering flexible working, within-sample. When sorting the sector work flexibilities in ascending order, we interestingly find a distinctive midpoint – 40% – above and below which work-flexibility levels are more homogenous.[30] As such, we separate our 11 sectors into two groups: six sectors with 'low' work flexibility (under 40% of firms in the sector) and five with 'high' work flexibility (above 40%, starting with 46% of firms in the sector). A *Highly Work-Flexible Sector* dummy is created taking the value of one if a company belongs to one of the five more work-flexible sectors and zero otherwise. Using the same settings in Table 4.3, we find that the *Highly Work-Flexible Sector* coefficient is positive and significant at the 1%-level whether controls are excluded or included (equal to 0.163 and 0.150,

[29] This is achieved by omitting the dummy for *Industry Work Flexibility Q1*, in line with the quartile analysis conducted by Lins et al. (2017).
[30] This median value is 40%, and the next sector flexibility observation is widely away at six percentage points higher.

respectively). This indicates that, on average, companies belonging to more work-flexible sectors saw great value protection during the COVID-19 crash, robust to the inclusion of controls for firm accounting and market-based measures.

4.5.3 Difference-in-Difference Analysis: Work flexibility during calm periods

We now endeavour to understand the relative importance of work flexibility during calm periods when flexible working, while potentially beneficial, is not a widespread necessity. We hypothesise that it only became relevant when investors collectively acted on expectations that work-flexible firms would face less disruption in the face of restrictions on social contact and travel. To explore this temporal relationship between work flexibility and stock performance, we conduct a difference-in-difference analysis including time and fixed effects. We compare the relationship between Work Flexibility and returns during our pandemic window to the calendar year preceding.[31]

We construct a panel of monthly returns starting in January 2019, one year before our observed crisis period, and ending in March 2020. We estimate the following model:

$$Return_{i,t} = b_0 + b_1 Work\ Flexibility_{i,2018/19} + b_2[Work\ Flexibility_{i,2018/19} \times Crisis_t] + b_3' \mathbf{X}_{i,t-1} + Time\ Dummies + Country\ Fixed\ Effects + Industry\ Fixed\ Effects + e_{i,t} \quad \textit{(Equation 4.3)}$$

$Return_{i,t}$ is the monthly raw return or abnormal return. Abnormal returns are calculated as CAPM-adjusted returns where model parameters are estimated over the five-year period ending in December 2018, using monthly data. $Work\ Flexibility_{i,2018/19}$ is a dummy for whether a firm offers flexible work arrangements as of 2018 for return observations in 2019,

[31] Our aim is to understand how the COVID-19 pandemic may have altered the importance of work flexibility. As such, we feel that the most comparable time period is only the year prior. Going back further may also capture how corporate culture with respect to work flexibility may have evolved, which would confuse our analysis. Extending the window further back would have also reduced the number of companies in our sample due to lack of data availability.

and as of 2019 for return observations in 2020. $Crisis_t$ is a dummy variable set to one for each of January, February, and March 2020, and $\mathbf{X}_{i,t-1}$ is a vector of controls, as employed in Table 4.2. Controls are lagged by one year for accounting variables and by one month for market variables. Four-factor loadings are re-estimated each month based on the previous 60 months' data. Time dummies are specified at the monthly level. Country and industry fixed effects are used to display, for reference, the coefficient of the *Work Flexibility* indicator itself. We also estimate a slightly different version of the model, dropping the work flexibility indicator and using firm fixed effects. These would absorb a firm's work flexibility (or lack thereof)[32] and control for all time-invariant firm-level omitted variables. In all models, standard errors are clustered at the firm level to account for the correlation of firm outcomes over time.[33] The coefficient on the interaction between work flexibility and the crisis captures the differential impact of work flexibility on stock performance during the first quarter of 2020, after controlling for firm characteristics, time-series patterns, and country and industry (or firm) effects.

The results are summarised in Table 4.4. Columns (1) and (2) show that work flexibility is negatively associated with stock performance outside of the COVID-19 crisis period, and significantly so when considering raw returns. This could be explained by the potential costs of work flexibility discussed in Section 4.3. However, the interaction term is positive under both return measurements, and significantly so in Column (1). This indicates that work flexibility added raw-return value during the COVID-19 crisis, unlike during the calm year preceding. A Wald-test performed on the positive sum of coefficients for each of Columns (1) and (2) is significant at the 1%-level and 10%-level, respectively. This reflects that investors were less likely to sell-off shares of companies whose employees were accustomed

[32] While this indicator is not taken as of a single time period in our panel, it does not change from 2018 to 2019 for the vast majority of the firms in our sample.
[33] Our results are not different when double clustering, by firm and time period, is used.

to flexible work arrangements when, and only when, working flexibly became necessary during the COVID-19 crash. Under our stricter models in Columns (3) and (4) using firm fixed effects, we are not able to observe *Work Flexibility*'s coefficient (because this is a firm-level indicator), but we find an expectedly positive and significant interaction term in line with our previous conclusion.

Table 4.4. Returns Surrounding the COVID-19 Crisis Period (Difference-in-Difference Analysis)

This table presents regression estimates of monthly raw and abnormal COVID-19 crisis-period returns from January 2019 to March 2020 as a function of a firm's work flexibility. The crisis is defined as January to March 2020. Abnormal returns are computed based on the market model using the K. French Developed Markets value-weighted index as the market proxy. Market model parameters are estimated over the five-year period ending in December 2018 using monthly data. *Work Flexibility* is a dummy variable obtained from Refinitiv's ESG database. *Crisis* is a dummy that takes the value of one for each of January, February, and March 2020. The control variables are as defined in Table 4.1; accounting variables are updated annually, and market variables are updated monthly. Four-factor loadings are estimated using the 60 monthly observations prior to each return observation. Industries are determined according to GICS. Country dummies are defined by country of headquarters. The control variables and returns are winsorised at the 1st and 99th percentiles. Standard errors are clustered at the firm level and are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Raw Return*	(2) *Abn. Return*	(3) *Raw Return*	(4) *Abn. Return*
Work Flexibility	-0.00306**	-0.000770		
	(0.00138)	(0.00134)		
Work Flexibility × *Crisis*	0.0192***	0.00509	0.0232***	0.00832*
	(0.00405)	(0.00365)	(0.00470)	(0.00427)
Firm characteristics	Yes	Yes	Yes	Yes
Four-factor loadings	Yes	Yes	Yes	Yes
Firm fixed effects	No	No	Yes	Yes
Country and Industry fixed effects	Yes	Yes	No	No
Time (monthly) fixed effects	Yes	Yes	Yes	Yes
Standard errors clustered by	Firm	Firm	Firm	Firm
N	36,790	36,906	36,790	36,906
Adjusted R-squared	0.340	0.072	0.365	0.122

4.5.4 Aggregate ESG Scores: Is work flexibility a proxy for general CSR performance?

Employees being provided with flexible working arrangements is one of the ways in which firms try to improve job satisfaction, which management hopes would in turn improve workforce productivity (e.g., CIPD, 2019). Under firm ESG mandates, such workforce-based initiatives typically fall under the Social Pillar, which captures overall firm commitment to various forms of societal development and engagement. Our baseline analysis in Section 4.5.1 showed that *Work Flexibility*'s coefficient is positive and significant at the 10%-level when controlling for firm characteristics as well as sector and country effects. However, it can be suggested that our single-dimensional indicator of work flexibility is acting as a proxy for more comprehensive ESG indicators. Following from that, one could presume that more aggregate ESG measures might be more successful in explaining return differential during the crash.

We test this by replacing our *Work Flexibility* dummy with multidimensional ESG-based indicators from the same Refinitiv database, using our firm-level analysis with full accounting, market-based, country, and sector controls as per Table 4.2, Column (2).[34] We first consider the *ESG Score* which reflects a firm's performance under all three pillars: Environmental, Social, and Governance, where each pillar is weighted with respect to its sector-level relevance (Refinitiv, 2021). We then look at the *Social Pillar Score* separately, as it allows us to focus on the aspect arguably most relevant during a crisis induced by health and safety concerns, significantly impacting society. We also break into the Social Pillar and separately consider the *Community Score* and *Workforce Score*. The former allows us

[34] We are satisfied with controlling for only sector effects at this stage since we do not see a strong intuitive link between granular, industry-level business nature and these multidimensional ESG measures. Therefore, to reduce loss in degrees of freedom, we use sector dummies and consider them sufficient controls. Our findings here are not different when using industry dummies.

to test whether performance was driven by a company's commitment to protecting public health, and the latter considers a company's overall effectiveness in providing job satisfaction as opposed to singling out the work flexibility feature.[35] The indicators are taken as of year-end 2019 where available and 2018 otherwise, maintaining the same sample used throughout the chapter.

Our findings are summarised in Table 4.5. The all-negative coefficients, nearly always significant, would indicate that *Work Flexibility* is unlikely to be acting as a proxy for such more general ESG measures. Our estimates would imply that better scores in any of the tested ESG-based measures worsened, as opposed to improved, a firm's COVID-19 crash returns. This is not strictly intuitive, but discussion of the relationship between aggregate ESG measures and COVID-19 performance is beyond the scope of this chapter about work flexibility.

[Table 4.5. is on the next page.]

[35] As defined within the Thomson Reuters Eikon platform through which the Refinitiv ESG database was accessed for this thesis.

Table 4.5. COVID-19 Sector-Controlled Returns and Alternative ESG Indicators

This table presents regression estimates of COVID-19 crisis-period returns on multidimensional ESG-based indicators and control variables. *Crisis-Period Returns* are measured as raw buy-and-hold returns over the period January to March 2020. The *ESG Score*, *Social Pillar Score*, *Community Score* and *Workforce Score* are obtained from the Refinitiv ESG database and rescaled to a value between 0 and 1 through division by 100. They are taken as of year-end 2019 where available and 2018 otherwise. Firm characteristics correspond to the control variables as defined in Table 4.1. Sectors are determined according to GICS. Country dummies are defined by country of headquarters. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Crisis-Period Return*	(2) *Crisis-Period Return*	(3) *Crisis-Period Return*	(4) *Crisis-Period Return*
ESG Score	-0.0756***			
	(0.000239)			
Social Pillar Score		-0.0413**		
		(0.000210)		
Community Score			-0.0216	
			(0.000156)	
Workforce Score				-0.0441**
				(0.000172)
Constant	-0.857***	-0.822***	-0.796***	-0.828***
	(0.0899)	(0.0888)	(0.0855)	(0.0885)
Full set of controls	Yes	Yes	Yes	Yes
Four-factor loadings	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
N	2,532	2,532	2,532	2,532
Adjusted R-squared	0.402	0.401	0.400	0.401

Our results are meaningful in showing that single-dimensional ESG estimates, like work flexibility, might prove more informative than aggregate ones in certain market conditions in which investors seek particular firm advantages. In conclusion, in agreement with Steffen (2021), we posit that ESG should not be seen as a 'one size fits all'-indicator. Instead, we reiterate that individual ESG elements can become incrementally relevant depending on prevailing economic conditions and the underlying drivers of market movements.

4.5.5 Within-Sector Benefits: Was work flexibility rewarded more in certain sectors?

Finally, we tackle one further analysis akin to subsampling, exploring whether the relationship between work flexibility and returns is strong within certain sectors. Having already identified that sectors with a higher proportion of flexible firms outperformed their counterparts during the crisis, we test for within-sector effects using interaction terms. We avoid subsampling to allow for direct coefficient comparison to ascertain relative return enhancements across sectors. In this test, we effectively separate the *Work Flexibility* dummy into 11 interaction dummies, one for each of our GICS sectors. The coefficient on each interaction allows us to observe the return impact of work flexibility specifically within that sector, on average. These interaction coefficients represent marginal effects relative to the sector's average performance, captured using sector dummies. Adding sector dummies allows us to control for exceptional outperformance or underperformance of certain sectors during the observation window. These would capture the impact of sector-level work flexibility on returns and ensure robustness against sector-specific downturns and booms which would have otherwise been captured by the interaction coefficients. In order to include all sector dummies and all interaction variables, the constant must be dropped to avoid perfect multicollinearity. This means that all sector dummies have negative coefficients, reflecting the fact that all sectors suffered a loss during the COVID-19 window on average.

This final analysis is somewhat 'experimental' in nature as it suffers from a key complication. As previously discussed in this thesis, the sector's flexibility has been found to be a significant determinant of the sector's performance. As such, when controlling for sector-wide effects, part of the work flexibility effect is captured. The complexity lent by this makes the results of this final analysis more difficult to interpret. However, we include

this analysis as a precursor to further work on within-sector effects of CSR, and also as an opportunity to reveal and discuss which subset of companies benefited most during this unusual time from having been accustomed to flexible working. The results are summarised in Table 4.6.

Table 4.6. Within-Sector Analysis of COVID-19 Returns and Work Flexibility

This table presents the impact of work flexibility on COVID-19 crisis-period returns within each of 11 sectors using interaction variables. The interactions variables are named as "[*Sector*] x *WF*" and are obtained by multiplying each sector dummy by the *Work Flexibility* (WF) indicator. For example, *Communication Services* x *WF* has the value one for a firm which belongs to the Communication Services sector *and* offered pre-2020 work flexibility; the interaction term is zero otherwise. *Raw Crisis-Period Returns* are measured as unadjusted buy-and-hold returns over the period January to March 2020. Work flexibility is based on data from the Refinitiv ESG database. Sectors are determined according to GICS. Each sector's dummy is reported under its interaction term. The control variables are as defined in Table 4.1. Country dummies are defined by country of headquarters. The control variables and returns are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Raw Crisis-Period Return*
Communication Services x *WF*	0.0426
	(0.0379)
Communication Services	-0.766***
	(0.0907)
Consumer Discretionary x *WF*	0.0065
	(0.0179)
Consumer Discretionary	-0.887***
	(0.0814)
Consumer Staples x *WF*	0.0378
	(0.0278)
Consumer Staples	-0.649***
	(0.0840)
Energy x *WF*	0.0079
	(0.0307)
Energy	-0.987***
	(0.0865)
Financials x *WF*	0.0047
	(0.0321)
Financials	-0.788***
	(0.0872)
Health Care x *WF*	0.0663**
	(0.0260)

Health Care	-0.663***
	(0.0850)
Industrials x *WF*	0.0123
	(0.0147)
Industrials	-0.808***
	(0.0814)
Information Technology x *WF*	0.0172
	(0.0209)
Information Technology	-0.750***
	(0.0829)
Materials x *WF*	-0.0219
	(0.0256)
Materials	-0.777***
	(0.0831)
Real Estate x *WF*	0.0133
	(0.0284)
Real Estate	-0.762***
	(0.0821)
Utilities x *WF*	-0.0238
	(0.0252)
Utilities	-0.600***
	(0.0835)
Constant	No
Full set of controls	Yes
Four-factor loadings	Yes
Country dummies	Yes
N	2,532
Adjusted R-squared	0.811

Only the Health Care interaction term is significant.[36] This would suggest that, even after accounting for the better performance of this sector during the COVID-19 recession – which we posit is partly due to the sector's high work flexibility – a Health Care firm having employees accustomed to flexible working protected the firm's performance during the first

[36] Standard practice when using interaction variables encourages the inclusion of the interaction's component variables. In this case, the component variables would be the *Work Flexibility* and sector dummies. However, if the *Work Flexibility* dummy is included, one of the sector interactions would have to be dropped. Since our findings indicate work flexibility being most value-enhancing in the Health Care sector, the *Health Care* x *WF* interaction could be dropped, making it the base effect. This setting turns all other interactions negative, indicating that work flexibility is less value-enhancing in all other sectors compared to the Health Care sector. We report only the specification with all interactions, including with Health Care, to demonstrate the sector decomposition of *Work Flexibility.*

quarter of 2020. The coefficient on *Health Care* x *WF* can be seen as a marginal effect relative to the sector's performance, captured using the Health Care sector dummy.

The inclusion of sector dummies is considered an effective way to control for unobserved sector effects; however, we cannot be certain that investor optimism with respect to Health Care firms in light of the pandemic is fully captured by the sector's dummy. We do note that Health Care's strong performance is reflected by the sector effects: *Health Care* has one of the greatest coefficients, with only *Utilities* and *Consumer Staples* at less negative values.

Work flexibility being rewarded most within the Health Care sector is in line with expectations, considering the driver of the recession. The COVID-19 pandemic, being a public health crisis, put unmatched pressure and focus on the Health Care sector to deliver abnormally high output levels in short periods of time. During the first half of 2020, there was increasing and urgent demand for Health Care firms to provide products and services as COVID-19 cases peaked and hospitalisations surged (McCabe et al., 2020). Kaye et al. (2020) find that lack of preparedness was a major contributor to the financial struggles faced by healthcare systems internationally. We build on this by suggesting that the Health Care companies most prepared to respond to the demands of these systems were advantaged during the COVID-19 crash. We posit that part of said preparedness was lent by a firm's employees being accustomed to flexible working. As such, work-flexible Health Care companies were better positioned to capture the sector's rising demand. In summary, having employees accustomed to flexible working was most important for Health Care companies and less so in other sectors which were less central to the crisis.

4.6 Conclusion

The COVID-19 pandemic gave rise to social distancing measures and travel restrictions which caused widescale disruption to the average work life. Our findings suggest that firms whose workforce were better prepared for such disruption, having been accustomed to work flexibility pre-2020, outperformed their counterparts during the first quarter of 2020, even when sector effects are accounted for. Work flexibility was not value-enhancing during 2019, the calm year leading up to the COVID-19 crisis. This suggests that the benefit of work flexibility to the firm is contingent on the presence of an exogenous shock which renders flexibility more valuable.

Interestingly, the inclusion of industry effects leaves work flexibility insignificant in explaining stock market returns. This indicates that industry-level performance during the COVID-19 crisis was closely linked to the industry's ability to operate a flexible workforce. We build on existing work by showing that this positive association is robust to the inclusion of firm accounting and market-based characteristics and the use of an alternative, more encompassing definition of work flexibility.

By using more aggregate ESG measures and finding quantitatively and qualitatively different results, we also show that work flexibility does not proxy for more general CSR engagement. Finally, a within-sector analysis revealed that work flexibility was most value-enhancing in the Health Care sector. This is intuitive considering that the recession was driven by a public health crisis, placing great emphasis on Health Care companies and sharply increasing demand for their products and services in a short period of time.

With more workers expecting to work from home in the 'new normal' (e.g., Finder UK, 2021), it is imperative to understand the true costs and benefits of work flexibility not only to employees but also to the firm. On the whole, we put forth that the value of work

flexibility is maximised in periods of high uncertainty when adaptability is in high demand, which is consistent with other forms of firm-level flexibility.

More broadly, our study sheds light on the value of preparedness. Our findings ultimately reveal that firms and industries with employees more prepared for the unforeseen COVID-19 pandemic, through having been accustomed to flexible working, outperformed those less prepared. Thinking ahead, while the COVID-19 crisis was unexpected, some ESG-related market-destructive events are more predictable. For example, climate change has increasingly become regarded as a matter of systemic importance (e.g., CFTC, 2020). Our findings on preparedness could serve as motivation for companies immunising themselves against market-wide events by proactively engaging in the appropriate CSR activities.

5 ENVIRONMENTAL RESPONSIBILITY AND FIRM PERFORMANCE DURING EXTREME FLOODING IN EUROPE

5.1 Abstract

As a key part of a firm's CSR, environmental responsibility can be captured by E-scores of widely used ESG measures. Utilised as a measure of 'greenness', it has recently gained further traction as a climate change risk management indicator. Using a familiar cross-sectional approach, we find that European firms with lower Environmental Scores outperformed their counterparts during the period of extreme flooding in July 2021. While this is not in line with a sentiment-based hypothesis which would have predicted higher demand for environmentally responsible stocks in the wake of a climate disaster, the findings could support a risk proxy and/or greenwashing hypothesis. Specifically, we suggest that E-scores are acting as a proxy for high environmental exposure: firms most in need of addressing their environmental risk would be most active in portraying good environmentalism, obtaining high E-scores. Our finding of a negative association between E-scores and financial performance may suggest that investors view score-targeted environmentalism as greenwashing and penalise it during extreme climate events.

5.2 Introduction

In recent years, environmental sustainability has rocketed to the top of governmental, corporate, and household agendas. In large part, this is because of climate change and extreme weather events becoming the "new norm" with people fearing for the long-term habitability of the planet (McGrath, 2021b). Examples of such events include heat waves in North America, droughts in South Africa, and record-breaking flash flooding in Western Europe – the last of which is the focus of this chapter.

From a financial perspective, climate change has been accompanied with increased regard for corporate environmental responsibility – a component of ESG criteria used by ethically-motivated investors in determining a firm's social responsibility. The business case for CSR has long been debated. However, the globally experienced COVID-19 pandemic is thought to have evolved the corporate ecosystem towards increased stakeholder capitalism (Hunt et al., 2020). Greater awareness of sustainability's role in economic development has even been considered a silver lining of the pandemic (Qiang et al., 2021).

In the EU, the increased focus on ESG factors since 2020 has also been attributed to significant regulatory developments, including the Sustainable Finance Disclosures Regulation (SFDR) and the Taxonomy Regulation (Norman et al., 2021). Increased regulation in this area is to be expected, given investors' increased appetite for sustainable business. As of December 2021, the UN Principles for Responsible Investment have been ratified by 2,775 institutional investors, representing over USD 121 trillion in assets under management (UNPRI, 2021). This is a marked increase from the USD 6.5 trillion committed when the Principles were first established in 2006, and from the pre-pandemic 2018 figure of USD 81.7 trillion (Eccles and Klimenko, 2019).

In the first half of 2021, several key steps were taken towards climate stabilisation. The European Commission put forth a plan on how the EU will reach climate neutrality by 2050 as part of the European Green Deal (European Commission, 2021). US-President Biden re-joined the Paris Climate Accords and made a commitment to make the US 'net-zero emissions' by 2050 (McGrath, 2021a), and the 26th United Nations Climate Change Conference (COP26) was highly anticipated by asset owners, regulators, and the public.

Amidst this heightened climate change coverage, multiple European countries experienced severe flooding in July 2021, causing numerous casualties and material economic loss (McGrath, 2021b). This study investigates equity market reactions to this climate disaster – the first significant one in the EU since the pandemic, occurring during a period of increased focus on environmentalism in the wake of new EU sustainability regulation and ahead of COP26. Using a sample of 370 European firms domiciled in the worst-affected countries, we explore the relationship between buy-and-hold equity market returns over the one-week flooding window and Environmental (E) Scores from the Refinitiv database.

Prior research, summarised below, suggests that green stocks – those that are more environmentally friendly – ought to outperform their counterparts during a climate disaster. If this is assumed to be true, a positive association in our study would suggest that investors consider Refinitiv's E-scores a reliable indicator of a firm being 'green'. Meanwhile, a negative association would suggest the contrary – perhaps indicating that high Refinitiv E-scores proxy for high environmental risk exposure instead. Our findings seem to support the latter hypothesis, as the Environmental Score coefficient in our cross-sectional regression is negative.

Looking to the E-score itself in explaining our findings is supported by a sizeable literature, summarised below, highlighting the disparity between ESG measures. It is possible that the

use of a different environmental indicator would reveal results different from previous work. Considering Refinitiv scores' high reliance on companies self-reporting, it is possible that the E-scores used in this analysis suffer from a degree of selective disclosure. Whether intentional on the part of the firms or not, we suggest that our results indicate investor belief that companies with higher E-scores are the ones with the most environmental concerns to address.

This chapter contributes to the literature in several key ways. First, to the best of our knowledge, this is the first study to investigate equity market reactions to the 2021 severe flooding window in Europe. This gives insight into new-age investor regard for corporate environmentalism following the paradigm shift brought about not only by the pandemic but also by new EU sustainability regulation. Second, our findings contribute to the responsibility-performance literature that continues to be divided. By focusing on an unexpected climate crisis, we contribute to the subset of the literature exploring the relative safety of green stocks during unexpected climate deterioration. Our findings do not appear to corroborate earlier suppositions that green stocks outperform brown ones during climate disasters. However, this may be due to our choice of alternative 'greenness' measurement – Refinitiv Environmental Scores. In adopting this different yet generally popular indicator in our study, we also contribute to the ESG measurement literature and confer with findings that different ESG indicators may deliver widely different results – strengthening the case for the development of more stable and standardised ESG scores.

The rest of this chapter is structured as follows. Section 5.3 summarises the literature and develops the hypothesis. Section 5.4 describes the data and methodology, and Section 5.5 presents the results of our testing. Finally, Section 5.6 concludes.

5.3 Related Literature and Motivation of Study

5.3.1 'Green' versus 'Brown' Stock Performance

As part of the broader literature on CSR, as well as on climate risk more recently, some studies focus only on the environmental aspect, as measured by the E-score and its components parts. For example, Garel and Petit-Romec (2021) find that stocks of firms with higher environmental responsibility, but not necessarily higher general corporate social responsibility, outperformed their counterparts during the COVID-19 pandemic.

The importance of the E-score may be on the rise given the planet's changing climate and increased environmental incidents. Climate change risk has been brought to the forefront of regulator agendas, with central banks globally recognising it as a systemic risk that governments must urgently address (Ramani, 2020). Investors, too, need to immunise their portfolios against climate risk. At the core of financial research on CSR is the question of whether a company can 'do well by doing good'. Similarly, a better understanding is needed of investor compensation – if any – for holding green stocks. Engle et al. (2020) hypothesise that environmentally friendly stocks would be more resilient to negative climate change news, providing an opportunity to hedge climate risk. The authors use MSCI and Sustainalytics E-scores to proxy for environmental friendliness and find that green stocks do have relatively higher returns than brown stocks when there is negative climate change news. It is worth noting that their sample was US-only and that they find key differences in the performance of the hedge portfolios depending on the provider of E-scores. They also note that they have few negative climate news observations, which subjects their findings to data mining concerns, as addressed in their study. This gives good reason to explore the performance of high-E-score companies using a different sample and score provider, as employed in our study. We also avoid having to rely on negative climate news over a time series by focusing on a single climate disaster window.

Pástor et al. (2021b) suggest that investors hold green stocks despite their lower expected return because they derive utility from having environmentally friendly portfolios and because green stocks hedge climate risk, as put forth by Engle et al. (2020). This is consistent with earlier findings that socially responsible investors are willing to sacrifice return for ethical preferences (e.g., Renneboog et al., 2011; Riedl and Smeets, 2017). In fact, the underperformance of green assets is well documented. In keeping with Hong and Kacperczyk (2009)'s prominent work on the outperformance of 'sin stocks', El Ghoul et al. (2011) find that lower-CSR firms, including those with worse environmental responsibility, obtain higher equity premia. Focusing solely on environmental concerns, Chava (2014) finds that firms with substantial emissions and climate change externalities also outperform their green counterparts.

However, in their empirical paper, Pástor et al. (2021a) find that green assets are yielding high returns in recent times. They put forth that this is caused by increased investor taste for environmental responsibility in light of peaking climate concerns, as theoretically introduced in their earlier paper on the same issue (Pástor et al., 2021b). Therein, the authors indicate that green assets would outperform brown ones if ESG concerns strengthen unexpectedly. Ardia et al. (2020) have also been able to find empirical support for Pástor et al. (2021b)'s theory using a different sample of U.S. stocks and a different measure of 'greenness' – namely, the CO_2-equivalent greenhouse gas indicator from Refinitiv's ESG database. Along with similar findings by Choi et al. (2020), prior work would indicate green stocks should gain value relative to brown stocks during unexpected climate deterioration, which could manifest as an unexpected extreme weather event (Carbon Brief, 2022).

This relates to broader literature that cites responsible assets earning superior gains during a specific window in which corporate responsibility is unexpectedly sought by investors (e.g., Albuquerque et al., 2020; Bae et al., 2021; Lins et al., 2017). Taking a similar

approach, we focus on the heavy flooding window in Western Europe and hypothesise that investor sentiment for green investments may have been exogenously altered in light of the unexpected climate event.

Unlike the climate studies thus far cited, we look at a European sample in a cross-sectional setting. We also rely on Refinitiv's overall Environmental Pillar Score as our measure of environmentalism. It is important to explore the different conclusions given rise to by using alternative ESG or greenness measures. This is due to the prominent disparity found between popular ESG indicators. Engle et al. (2020) discuss how the climate risk hedge portfolio's success is highly influenced by whether MSCI or Sustainalytics scores are used in its construction. More broadly, Jacobs and Levy (2022) caution that ESG scores may not necessarily align with each other nor with true firm ESG performance. Christensen et al. (2022) also show that ESG ratings, particularly in relation to ESG outcomes, can be vastly different between providers.

The concern raised by this is whether the relationship between performance and greenness can truly be captured by E-scores, or whether these scores may be capturing something else entirely. A growing literature focuses on ESG data greenwashing, whereby companies report information on environmental practices selectively to portray an image misaligned with the company's true environmentalism (In and Schumacher, 2021). Firm managers have also been found to overstate the company's environmental performance as a way of deflecting scrutiny over earnings underperformance (Hail et al., 2021). It is worth noting that most ESG scores, including Refinitiv's used in this study, heavily rely on firms self-reporting their CSR activities which increases the potential for greenwashing (Yu et al., 2020). Concerningly, Tang et al. (2021) also find that firms whose owners hold a considerable stake in the ESG provider itself receive higher ESG scores, despite having poorer ESG outcomes.

5.3.2 Hypothesis Development

With this in mind, we develop our hypothesis considering that a positive, negative, or non-existent relationship between E-scores and flooding-window performance may be found. A negative association could be justified if investors perceive higher E-scores to be greenwashed. We state our hypothesis below.

We decide to test the association between environmental responsibility and firm performance during the period of extreme flooding in Europe. Both a positive and negative association can be supported intuitively.

Hypothesis 1a: There is a positive association between environmental responsibility and firm performance during a climate incident.

This hypothesis is founded on sentiment investing. In light of an extreme weather event, investors would flock towards more environmentally responsible firms that they believe are helping the environment or, at least, avoiding harm to it. From a risk perspective, it is possible that investors believe high-E stocks to be immunised against unexpected climate deterioration, leading them to outperform during a climate disaster.

Hypothesis 1b.: There is a negative association between environmental responsibility and firm performance during a climate incident.

If this hypothesis is supported, we suggest it is because an assumption is being made that firms with higher E-scores are the ones more exposed to climate risk and hence taking the most steps towards ensuring good environmental practice. A sinister take on this would align to a greenwashing hypothesis: brown firms use score-increasing environmentalism to misrepresent themselves as green. If investors believed this to be the case, they would move away from high-E stocks. Regardless of firm intentions, this hypothesis associates higher E-scores with riskier, poorer environmentalism.

If we are unable to reject the null hypothesis of no association, we can infer that European market participants do not use environmental responsibility to guide their short-term investment decision making around climate event windows. We describe our data for testing this hypothesis in the next section.

5.4 Data and Methodology

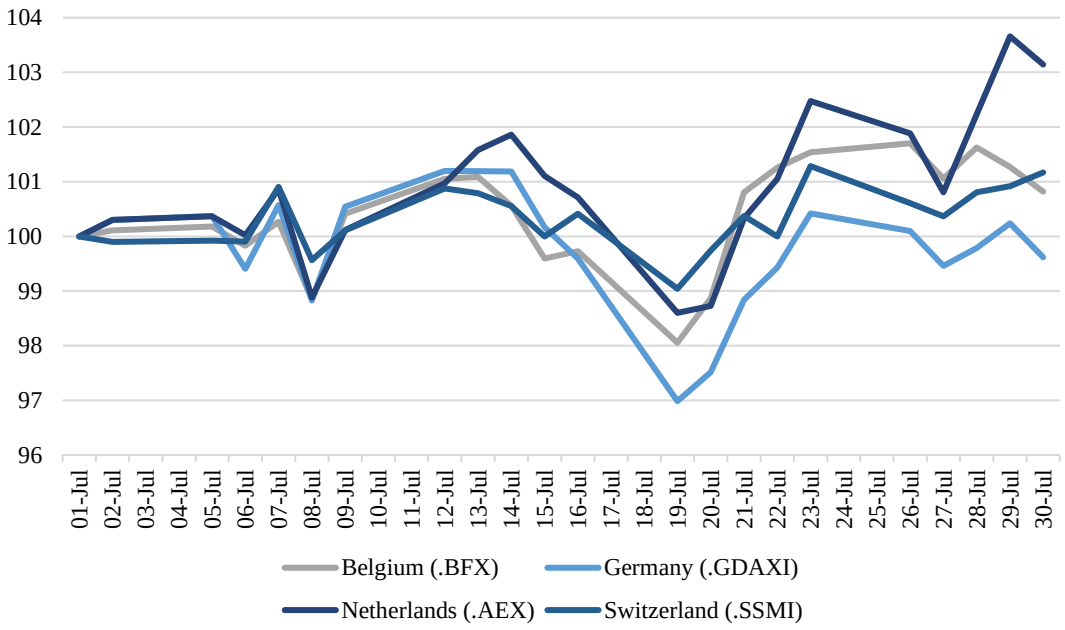
Our sample comprises companies headquartered in Germany, Belgium, Netherlands, and Switzerland. Germany is considered the worst-hit country, with the floods claiming close to 180 lives (Schmidt et al., 2021). The same source indicates that Belgium had the next highest death toll, at 38 lives lost. Data from Switzerland's Federal Office for the Environment recorded river discharges that mark these summer floods as a 'once in 300 years'-event (JBA Risk, 2021). The Maas River water outflows also affected parts of the Netherlands, with the Dutch Infrastructure Ministry citing 100-year water level records being broken and the need for army intervention (Dutch News, 2021). Serving as the dependent variable in our cross-sectional analysis, buy-and-hold returns are calculated over the window from 9 to 19 July 2021.[37] This window is based on data from NASA's Precipitation Processing System indicating peak rainfall in the affected countries during this time (JBA Risk, 2021). We observe that this climate disaster was reflected in the performance of the relevant countries' main stock indices during this window, as can be seen in Figure 5.1.

Figure 5.1. Relevant European Markets' Equity Indices, 1 July 2021 = 100

The figure shows the value of a nominal 100 invested in each of the main equity market indices in Belgium (BFX), the Netherlands (AEX), Germany (DAX), and Switzerland (SSMI) – the countries

[37] Peak rainfall was experienced from 11 to 18 July 2021 (JBA Risk, 2021). As these two dates fall on Saturdays, a complete understanding of the market's change from just before the worst of the flooding window to just after must be 9 to 19 July 2021 (beginning and ending on trading days).

worst affected by the 2021 summer floods in Europe – over the period 1 July 2021 to 30 July 2021. This figure shows the high synchronicity between the main equity indices of these European countries, as well as notable market downturns between 9 and 19 July, taken as our flooding window for return-calculation purposes. NASA data cited in JBA Risk (2021) places peak rainfall in this time window. Price data was obtained from Refinitiv via Thomson Reuters Eikon.



To measure environmental responsibility, we use the Environmental Pillar Score from Refinitiv's ESG database, *Environmental Score* hereafter. Using a proprietary algorithm, Refinitiv scores a company's environmentalism against three main themes: Resource Use, Emissions, and Innovations. The score factors in whether a company has policies in place relating to environmental friendliness, initiatives, and reporting, as well as figures on emissions and energy use (Refinitiv, 2021). We scale this score to a value between 0 and 1 (simply by dividing the Refinitiv percentile score by 100).

We select a set of controls to be used in our regression based on similar cross-sectional work (e.g., Lins et al. (2017)). These variables are a combination of accounting figures, firm characteristics, and market indicators. The accounting variables are measured at the end of 2020, the nearest year-end prior to the flooding window. *Long-Term Debt*, *Short-Term Debt*, *Cash Holdings* and *Profitability* are included as a ratio to Total Assets. Equity *Book-to-*

Market and a dummy for *Negative B/M* firms are included.[38] *Momentum* is measured as the firm's holding-period return over the 12 months preceding the crisis period, and *Volatility* is the standard deviation of weekly returns over the same period. All non-dummy variables are winsorised at the 1st and 99th percentiles to mitigate the effect of outliers. All financial and accounting data is obtained from Refinitiv via Thomson Reuters Eikon (unless otherwise stated). We exclude companies for which the required accounting data is unavailable. This yields a final sample of 370 companies. Table 5.1 summarises the statistics for our dependent variable of buy-and-hold *Flooding-Window Return*, our key independent variable of *Environmental Score*, as well as all controls.

In Panel A, we note that firms in our sample experienced negative returns during our flooding-window. This was a loss of 3% during the single week of flooding – an economically significant decline. Panel B shows no correlation between the *Environmental Score* and *Flooding-Window Return*. This is not in line with some earlier work (e.g., Garel and Petit-Romec, 2021) which would have predicted a positive association, especially if the flooding period is considered to be a crisis. However, we recognise that a regression analysis controlling for firm, accounting, and sector differences would reveal a more robust association. This is presented in Section 5.5.

[38] As previously mentioned, this corrects for the behaviour exhibited by negative book-to-market firms which are likely to be in distress and hence act more like high book-to-market firms (Fama and French, 1992).

Table 5.1. Descriptive Statistics: Flooding Window and Environmental Scores

The sample consists of 370 firms headquartered in Germany, Belgium, Switzerland, or the Netherlands with Environmental Pillar Scores from the Refinitiv ESG database as of year-end 2020. *Environmental Score* is Refinitiv's percentile Environmental Pillar Score as of year-end 2020, divided by 100. *Flooding-Window Return* is the raw buy-and-hold return computed over the period 9 to 19 July 2021. Accounting data is based on the last quarter of 2020 and is obtained from Refinitiv via Thomson Reuters Eikon. *Market Capitalisation* is in millions of euros. *Long-Term Debt* is computed as long-term debt divided by total assets. *Short-Term Debt* is computed as debt in current liabilities divided by total assets. *Profitability* is computed as operating income divided by total assets. *Cash Holdings* is computed as cash and marketable securities divided by total assets. *Book-to-Market* is computed as book value of equity divided by market value of equity. *Negative B/M* is a dummy variable set to one when the book-to-market ratio is negative and zero otherwise. *Momentum* is the buy-and-hold return from 9 July 2020 to 8 July 2021, the year leading up to the flooding window. *Volatility* is the standard deviation of weekly returns over the same year preceding the flooding window. All variables are winsorised at the 1st and 99th percentiles.

Panel A: Summary Statistics, *N* = 370					
	Mean	*SD (Std Dev)*	25th perc.	Median	75th perc.
Environmental Score	0.46	0.28	0.24	0.48	0.69
Flooding-Window Return	-0.03	0.04	-0.06	-0.02	0.00
Market Capitalisation (€m)	9,658	21,417	697	2,079	7,290
Long-Term Debt	0.22	0.16	0.08	0.20	0.32
Short-Term Debt	0.05	0.07	0.01	0.03	0.07
Profitability	0.03	0.11	0.00	0.04	0.08
Cash Holdings	0.17	0.17	0.06	0.13	0.23
Book-to-Market	0.61	0.65	0.20	0.44	0.82
Negative B/M	0.01	0.12	0.00	0.00	0.00
Momentum	0.45	0.50	0.10	0.32	0.66
Volatility	0.05	0.02	0.03	0.04	0.05

(Continued…)

Panel B: Correlation Matrix, *N* = 370										
	Env. Score	*F.W. Return*	*Ln(Mkt Cap)*	*LT Debt*	*ST Debt*	*Profit.*	*Cash Hold.*	*B/M*	*Neg. B/M*	*Mom.*
Flooding-Window Return	0.00									
Ln(Mkt Cap)	0.43	0.05								
LT Debt	0.09	-0.13	0.04							
ST Debt	-0.01	0.01	-0.01	0.07						
Profitability	0.21	0.31	0.13	-0.13	0.04					
Cash H.	-0.28	-0.14	-0.09	-0.21	-0.13	-0.34				
B/M	-0.06	-0.19	-0.11	-0.01	-0.01	-0.28	-0.23			
Negative B/M	-0.06	0.02	-0.05	0.31	-0.05	-0.16	0.17	-0.11		
Momentum	0.04	0.06	0.00	-0.02	0.04	0.00	-0.02	0.00	0.05	
Volatility	0.00	0.03	0.02	0.04	0.07	0.02	-0.06	0.00	0.03	0.41

5.5 Flooding-Window Returns and Environmental Responsibility

The dependent variable is *Flooding-Window Return*, and we attempt to explain differences therein using the *Environmental Score* along with control variables, the firm's market-model beta, and country and sector dummies. For this test, the regression equation is as follows, and our results are reported in Table 5.2:

$$\textit{Flooding-Window Return}_i = \alpha + \beta_1 \textit{Environmental Score}_i + \beta_2 \textit{Ln(Market Cap)}_i + \beta_3 \textit{Long-Term Debt}_i + \beta_4 \textit{Short-Term Debt}_i + \beta_5 \textit{Profitability}_i + \beta_6 \textit{Cash Holdings}_i + \beta_7 \textit{Book-to-Market}_i + \beta_8 \textit{Negative B/M}_i + \beta_9 \textit{Momentum}_i + \beta_{10} \textit{Volatility}_i + \textit{Market-Model Beta} + \textit{Sector Dummies} + \textit{Country Dummies} + e_i \quad \textit{(Equation 5.1)}$$

The coefficient on *Environmental Score* is negative and significant in Column (1). This would indicate that firms with higher environmental responsibility experienced worse returns during the flooding window in Europe. This is also economically significant, where one standard deviation increase in E-score is associated with an underperformance of 0.46% during the flooding week. This is unexpected from a perspective of sentiment-driven investing, as investors flocking towards good environmentalism in the face of a climate incident would have yielded a positive relationship between *Environmental Score* and returns. The results may be explained by investors viewing a firm's pursuit of higher E-scores as either the need for environmental exposure mitigation or plain greenwashing, in line with Hypothesis 1b. In other words, firms that are more vulnerable to climate change may be over-engaging in 'softer' E-increasing activities and earn lower returns during a climate event.

Table 5.2. Flooding-Window Returns and Environmental Scores

This table presents regression estimates of equity returns during a window of extreme flooding on firm environmental responsibility scores for a sample of 370 European companies headquartered in the worst affected countries: Germany, Belgium, Switzerland, and the Netherlands. *Flooding-Window Return* is measured as the buy-and-hold return over the period 9 to 19 July 2021. We divide the Environmental Pillar Score from Refinitiv's ESG database by 100 to obtain our *Environmental Score* which ranges from 0 to 1. The control variables are as defined in Table 5.1. Sectors are determined according to GICS. Countries are determined by country of headquarters. The market-model beta is retrieved from Refinitiv via Thomson Reuters Eikon as the firm's 3-year beta up to the flooding window. All non-dummy variables are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1)	(2)
	Flooding-Window Return	*Flooding-Window Return*
Environmental Score	-0.0163*	-0.0161
	(0.00987)	(0.00987)
Ln(Market Cap)	0.00171	0.00168
	(0.00172)	(0.00172)
Long-Term Debt	-0.0377**	-0.0379**
	(0.0171)	(0.0172)
Short-Term Debt	-0.00961	-0.0112
	(0.0300)	(0.0300)
Profitability	0.0706***	0.0707***
	(0.0268)	(0.0269)
Cash Holdings	-0.0463**	-0.0472**
	(0.0193)	(0.0196)
Book-to-Market	-0.0104*	-0.0104*
	(0.00602)	(0.00598)
Negative B/M	0.0343*	0.0343*
	(0.0204)	(0.0205)
Momentum	0.00616	0.00570
	(0.00504)	(0.00501)
Volatility	0.0395	0.00163
	(0.106)	(0.125)
Constant	-0.0476	-0.0473
	(0.0362)	(0.0361)
Market-model beta	No	Yes
Sector and country dummies	Yes	Yes
N	370	370
Adjusted R-squared	0.149	0.148

In an attempt to use an environmental measure less susceptible to greenwashing, we replace the all-round *Environmental Score* with environmental indicators relating specifically to whether a firm exceeds a Refinitiv-set threshold for taking on environmental initiatives. We suggest that these measures may be less subject to greenwashing as environmental initiatives and expenditures are easier to directly verify from company annual reports and financial statements.

Namely, these variables are the *Environmental Investments Initiatives* and *Environmental Expenditures Investments* dummies from Refinitiv's ESG database. The two variables are defined very similarly by Refinitiv, so we employ both for robustness. Refinitiv provides the installation of cleaner technologies as an example of environmental initiatives. Both variable descriptions indicate that the dummy takes the value of one if the company "reports on making proactive environmental investments or expenditures to reduce future risks or increase future opportunities", as defined by Refinitiv (accessed via the Thomson Reuters Eikon database). However, the latter variable's definition also explicitly captures companies that report their environmental expenditures. The former variable is available for all companies in our sample while the latter is only available for 300 firms. Where both variables are available, the correlation between them is 82.3%, so we are satisfied that they capture very similar criteria.

Though we are unable to evaluate the success of these initiatives, we suggest that these indicators are less clouded by 'softer' E-score components that a firm could address without exercising tangible environmental responsibility. If a positive coefficient is observed on these dummies when they are used as the measure of environmental responsibility in the regression above, it is possible that investors trust more tangible environmental investments over aggregate E-scores in ascertaining a firm's environmental responsibility. On the other hand, one could argue that the firms spending the most on environmental initiatives are

perhaps the ones most exposed to environmental risk. If this is true, we would expect a negative coefficient on these dummies. The model from the previous analysis (including the market-model beta) is repeated with each of the dummies in place of *Environmental Score* to explore this. The results are reported in Table 5.3. While the coefficients on these variables are positive, they are not significant. We therefore suggest that this variable is more likely to be regarded as true environmental responsibility rather than a proxy for environmental risk, but we cannot say so conclusively.

Table 5.3. Flooding Window Returns and Environmental Investments Dummies

This table presents regression estimates of equity returns during a window of extreme flooding on a dummy variable indicating whether a firm invests in environmental initiatives. The sample is of 370 (Column 1) and 300 (Column 2) European companies headquartered in the countries worst affected by the flooding: Germany, Belgium, Switzerland, and the Netherlands. *Flooding-Window Return* is measured as the buy-and-hold return over the period 9 to 19 July 2021. *Environmental Investments Initiatives* (Column 1) and *Environmental Expenditures Investments* (Column 2) are dummies that take the value of one if a company "reports on making proactive environmental investments or expenditures to reduce future risks or increase future opportunities" (Refinitiv via Eikon), and zero otherwise. The control variables are as defined in Table 5.1. Sectors are determined according to GICS. Countries are determined by country of headquarters. The market-model beta is retrieved from Refinitiv via Thomson Reuters Eikon as the firm's 3-year beta up to the flooding window. All non-dummy variables are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Flooding-Window Return*	(2) *Flooding-Window Return*
Environmental Investments Initiatives	0.00177	
	(0.00516)	
Environmental Expenditures Investments		0.00379
		(0.00520)
Ln(Market Cap)	-2.88e-05	-0.000909
	(0.00135)	(0.00143)
Long-Term Debt	-0.0386**	-0.0358*
	(0.0169)	(0.0200)
Short-Term Debt	-0.0125	0.000233
	(0.0287)	(0.0284)
Profitability	0.0718***	0.0927***
	(0.0267)	(0.0330)
Cash Holdings	-0.0426**	-0.0564**

	(0.0196)	(0.0259)
Book-to-Market	-0.0110*	-0.00997
	(0.00597)	(0.00697)
Negative B/M	0.0331	0.0419
	(0.0204)	(0.0258)
Momentum	0.00545	0.00801
	(0.00506)	(0.00601)
Volatility	-0.00241	-0.116
	(0.127)	(0.135)
Constant	-0.0206	0.0147
	(0.0316)	(0.0345)
Market-model beta	Yes	Yes
Sector and country dummies	Yes	Yes
N	370	300
Adjusted R-squared	0.141	0.129

For robustness, we also consider a handful of other ESG measures from the Refinitiv database. We replace *Environmental Score* from our baseline analysis with each of the *Social Score*, *Governance Score*, aggregate *ESG Score*, and *ESG Controversy Avoidance Score*.[39] While Refinitiv also provides a score for how well a firm avoids environmental controversies specifically, this variable is a 'full mark' (i.e., no penalty) for the vast majority of the firms in our sample, indicating no environmental controversies. As such, it is not possible to run a regression using this variable.

The findings of these tests are reported in Table 5.4. The only significant result is for the *ESG Controversy Avoidance Score*, which has a positive coefficient. This means that firms with fewer ESG controversies outperformed their counterparts during the flooding window. Though intuitive, this result should be looked at cautiously since over 85% of the firms in

[39] These are the "Social Pillar Score", "Governance Pillar Score", "ESG Score", and "ESG Controversies Score" from the Refinitiv database, divided by 100 to create a scale of 0 to 1.

the sample have a 'full mark' on this, with the average *ESG Controversy Avoidance Score* at 0.92 (out of a maximum of one).

Table 5.4. Flooding Window Returns and ESG Measures

This table presents regression estimates of equity returns during a window of extreme flooding on various ESG measures. The sample is of European companies headquartered in the countries worst affected by the flooding: Germany, Belgium, Switzerland, and the Netherlands. *Flooding-Window Return* is measured as the buy-and-hold return over the period 09 July 2021 to 19 July 2021. *Social Score*, *Governance Score*, *ESG Score*, and *ESG Controversy Avoidance Score* are obtained from the Refinitiv database as of 2020. The control variables are as defined in Table 5.1. Sectors are determined according to GICS. Countries are determined by country of headquarters. The market-model beta is retrieved from Refinitiv via Thomson Reuters Eikon as the firm's 3-year beta up to the flooding window. All non-dummy variables are winsorised at the 1st and 99th percentiles. Heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Flooding-Window Return*	(2) *Flooding-Window Return*	(3) *Flooding-Window Return*	(4) *Flooding-Window Return*
Social Score	-0.00937			
	(0.0122)			
Governance Score		-0.0149		
		(0.0112)		
ESG Score			-0.0160	
			(0.0140)	
ESG Controversy Avoidance Score				0.0333***
				(0.00933)
Constant	-0.0347	-0.0390	-0.0432	-0.0953**
	(0.0352)	(0.0348)	(0.0374)	(0.0386)
Controls	Yes	Yes	Yes	Yes
Market-model beta	Yes	Yes	Yes	Yes
Sector and country dummies	Yes	Yes	Yes	Yes
N	347	347	347	358
Adjusted R-squared	0.172	0.175	0.174	0.192

Finally, we consider whether the credibility of the E-score increases when the company has invested in environmental initiatives or has good social and governance scores. We do this by estimating the coefficient on each of four interaction terms. *Initiatives* x *E-Score* and *Expenditures* x *E-Score* multiply the relevant dummy (*Environmental Investments Initiatives* and *Environmental Expenditures Investments*, respectively) by the *Environmental Score*. This allows us to investigate the correlation between the *Environmental Score* and performance during the flooding window in the subset of firms that have made tangible environmental investments. We also compute the dummy variables *High-SG(Median) and High-SG(TopQ)*, which take the value of one if the firm's average of the *Social Score* and *Governance Score* (*SG*) is above the median or in the top within-sample quartile, respectively. We then interact each of these dummies with the *Environmental Score*. The coefficient on these variables will allow us to see how the *Environmental Score* is associated with performance when the company also has good social and governance metrics.

We hypothesise that the companies in these subsets may be more sincere in their environmental efforts, possibly reverting the relationship between the E-score and performance during the flooding window to a positive one. We summarise these findings in Table 5.5. The *Environmental Score* is still negatively associated with returns, and the marginal effect for all subsets is also negative, though insignificant. These results do not lend support to there being a positive relationship between environmental scores and performance for the aforementioned subsets of firms during the flooding window.

Table 5.5. Flooding Window Returns and Interaction Variables

This table presents regression estimates of equity returns during a window of extreme flooding on interaction variables that indicate the relationship between the *Environmental Score* (*E-Score*) and returns for subsets of firms. The sample is of European companies headquartered in the countries worst affected by the flooding: Germany, Belgium, Switzerland, and the Netherlands. *Flooding-Window Return* is measured as the buy-and-hold return over the period 9 to 19 July 2021. *Initiatives* x *E-Score* is equal to the dummy *Environmental Investments Initiatives* multiplied by the *Environmental Score*. *Expenditures* x *E-Score* is equal to the dummy *Environmental Expenditures Investments* multiplied by the *Environmental Score*. *High-SG(Median)* x *E-Score* and *High-SG(TopQ)* x *E-Score* are interactions of the *Environmental Score* with a dummy equal to one if the firm's average of the *Social Score* and *Governance Score* is greater than the median and top-quartile within-sample, respectively. The standalone variables are also included. The control variables are as defined in Table 5.1. Sectors are determined according to GICS. Countries are determined by country of headquarters. The market-model beta is retrieved from Refinitiv via Thomson Reuters Eikon as the firm's 3-year beta up to the flooding window. All non-dummy variables are winsorised at the 1st and 99th percentiles. White (1980) heteroscedasticity-consistent standard errors are presented in parentheses. ***, **, and * indicate that the parameter estimate is significantly different from zero at the 1%, 5%, and 10% level, respectively.

	(1) *Flooding-Window Return*	(2) *Flooding-Window Return*	(3) *Flooding-Window Return*	(4) *Flooding-Window Return*
Initiatives x *E-Score*	-0.0343			
	(0.0219)			
Expenditures x *E-Score*		-0.00993		
		(0.0218)		
High-SG(Median) x *E-Score*			-0.0232	
			(0.0193)	
High-SG(TopQ) x *E-Score*				-0.0338
				(0.0211)
Environmental Investments Initiatives	0.0260*			
	(0.0153)			
Environmental Expenditures Investments		0.0135		
		(0.0147)		
High-SG(Median)			0.00851	
			(0.0109)	
High-SG(TopQ)				-0.0338
				(0.0211)
Environmental Score	-0.0142	-0.0221*	-0.00421	-0.00944
	(0.0106)	(0.0120)	(0.0122)	(0.0112)
Constant	-0.0507	-0.0235	-0.0659	-0.0643

	(0.0366)	(0.0404)	(0.0403)	(0.0391)
Controls	Yes	Yes	Yes	Yes
Market-model beta	Yes	Yes	Yes	Yes
Sector and country dummies	Yes	Yes	Yes	Yes
N	370	300	370	370
Adjusted R-squared	0.149	0.136	0.147	0.151

5.6 Conclusion

We test the relationship between Refinitiv E-scores and stock market performance for 370 European firms during flash floods in Europe, taken to be an unexpected climate change event. While prior research predicts the outperformance of green stocks during such a window, we find the opposite. Firms with high E-scores underperformed their counterparts by an economically significant margin. There are several possible reasons for this. One may be that the flash floods were not considered enough of an unexpected climate occurrence to trigger a skew towards greenness. However, the numerous casualties and significant economic losses suffered as a result of this extreme weather event would suggest otherwise.

Alternatively, we discuss that this unexpected finding may be due to investor perception of the environmental score itself. It is possible that investors active in the European markets we study associate higher E-scores with increased environmental exposure instead of mitigated climate risk. This would not necessarily negate previous findings which examine different markets using a different methodology, and importantly, use a greenness indicator that is different to ours. ESG measurement research has already documented that ESG scores from the major providers can greatly disagree on the level of a firm's environmental responsibility.

By limiting our sample to firms headquartered in the worst affected countries, we assume some homogeneity in the impact of the flooding on firm operations. However, extensions

to this work could also consider firm proximity to the worst affected areas within these countries. This is to reflect the findings of existing studies on stock market responses to extreme weather events which show that returns may only be impacted if the company can be considered 'local' to the disaster (Bourdeau-Brien and Kryzanowski, 2017; Donadelli et al., 2020). This can be facilitated by making use of geocoded natural disaster datasets in conjunction with more specific information about company geographical information, including its city of headquarters and the number of branches it has in different regions (Venturini, 2022). We also recognise that the companies captured by our sample may not be the ones worst affected by the floods on account of their large size and likely geographically diversified earnings. Small to medium-sized enterprises would likely be more vulnerable to natural disasters in general (e.g., Hong et al., 2019; Venturini, 2022). However, since our study is focussed on the association between financial performance and environmental scores, using a sample of large, listed firms for which ESG data is available remains reasonable. It would also be interesting to investigate whether lenders, not only shareholders, consider a firm's environmental performance in the wake of a major climate incident. While less likely to be driven by sentiment, lenders may regard firms with higher E-scores to be better protected from climate change and offer them more favourable terms. The effects of this would likely be observed over longer-term periods.

Our study serves as fertile ground for further work on the link between environmental friendliness and firm performance, with a need for critical evaluation of greenness measures. Prior work showing green stock outperformance during climate deterioration is largely US-focused. Our study extends the conversation to other regions, taking something of a case study approach by looking at the four European countries most affected by the 2021 flash floods. Given our findings, it is possible that climate risk-hedged portfolios in the style of Engle et al. (2020) may not prove successful for European market portfolios, and this would

be worth further exploration. Comparing how different markets react to climate change events may help us glean valuable lessons about the effects of differing regulations, particularly with respect to sustainability disclosures which remain quite disparate between the US and the EU. Understanding how climate risk is priced in equity markets during calm periods as well as periods of climate change deterioration is of value to investors, regulators, and environmental policymakers alike. This area warrants further exploration, particularly given the systemic risk climate change presents not only to economies but to the planet itself.

6 CONCLUSIONS

This thesis sets out to explore the relationship between CSR and CFP during equity market downturns. Our findings contribute to the abundant, yet inconclusive, literature on this link. It further scrutinises the credibility of the ESG scores widely used in research. The empirical work presented looks at each of the following equity market downturns: the Subprime Crisis in 2008/9, the Sovereign Debt Crisis in 2011, the COVID-19 stock market crash in early 2020, and the short-term European market decline in July 2021 that coincided with unprecedented flash flooding. For each of these periods, the relationship between CSR and equity market performance was investigated. For each of the studies, a select sample of publicly listed corporations from across sectors and countries was utilised. Further, depending on the nature of the downturn – including the cause of the crisis, prevailing market sentiments, and contemporaneous issues of social importance – the most relevant CSR dimensions were isolated and explored.

We find that the relationship between CSR, measured using aggregate scores or particular identifiers, and financial performance is difficult to summarise or generalise. A positive and significant association is found between ES-scores and market returns for European countries during the Subprime Crisis. However, the opposite relationship – negative and significant – is found during the Sovereign Debt Crisis a mere two years later. We put forth that this may be because of how strongly high ES-scores were connected to year-ahead ESG controversies. Next, we find that companies whose employees were accustomed to flexible working prior to the COVID-19 pandemic earned superior market returns during the pandemic-driven stock market crash in 2020. However, this appears to be primarily an effect of business nature, where a within-sector market reward for flexible working is only observed among Health Care companies. Finally, during the flash floods in Europe, a negative association is found between environmental scores and equity market performance.

This negative link disappears when ‘greenness’ is instead measured using dummy variables for whether a firm makes environmental investments – a more transparent and verifiable indicator.

A limitation of this work lies in the sensitivity of the results to the time periods considered. It is not always straightforward to determine the start and end date of an equity market downturn. We have relied on existing work, news, and observation of price momentum in deciding this for our analyses. We also perform robustness tests for reassurance that our results are not meaningfully different if a different sample period is chosen. However, being able to identify exactly when investor preferences changed would add credibility to the relationships between CSR and return being investigated.

Another limitation involving time is the lag that inherently exists between a firm undertaking CSR activities and the CSR scores reflecting this being made available to investors. Similarly to the availability of audited annual financial statements, there is some time between the calendar year-end and the release of ESG scores for that past year. As such, for an investor to know how well a company scored on responsibility metrics in, say, 2020, the investor would have to wait until the second or third quarter of 2021. This can be observed directly from the Refinitiv database which was used to obtain ESG ratings in this thesis. This means that potentially outdated scores are being used for time-sensitive analyses. While this is an unavoidable issue, we consider two mitigations. First, where feasible, we use ESG measures from different years for robustness to show that the results do not often change meaningfully. This is in part because a firm’s ESG score tends to be quite stable over time. Second, using lagged CSR measures allows us to be further reassured that investors are truly aware of the company’s most recently knowable CSR rating, and that scores too close to a crisis window are avoided since these could be somewhat influenced by the turmoil ahead (Lins et al., 2017).

We also note that our analyses are specific to developed markets, and we do not posit that our findings can be generalised to emerging markets. There is evidence that CSR disclosure is greater in developed markets (e.g., Bhatia and Makkar, 2020) as firms in developing countries do not receive as much public pressure to disclosure CSR initiatives (Ali et al., 2017). Therefore, our suppositions that investors resist the pressure to sell high-CSR stocks during a market downturn on account of a preference for trust, social responsiveness, or environmental friendliness may not hold outside of the countries considered in this thesis. However, if responsibility and sustainability are to be achieved at the global level, considerable emphasis must be placed on emerging markets where there is great room for improvement. According to the Global Rights Index 2021, the worst region in the world for working people is the Middle East and North Africa, with the top 10 worst countries also all developing (ITUC, 2021). Moreover, the majority of the countries ranked 'Very Low' on the Climate Change Performance Index 2022 are emerging economies (CCPI, 2022). This provides fertile ground for future work that explores drivers of CSR in emerging markets and how better CSR can be encouraged in them.

Furthermore, our empirical work uses a single source of ESG data for all analyses – the Refinitiv database, accessed via Thomson Reuters. While this has been increasingly used for research globally, there is evidence that vastly different conclusions can be drawn when scores from different providers are used (e.g., Berg et al., 2019; Jacobs and Levy, 2022). Unfortunately, the results of our analyses could not be cross-checked against ESG data from other providers, like MSCI, due to data access limitations. This brings to light areas for future work. The findings of these analyses could be compared to those using alternative ESG measures. While obtaining similar outcomes would lend credibility to the results presented in this thesis, differences could be highly informative. When empirical relationships do not align to theorised expectations, the limitation could lie in the credibility

of the ESG measure itself. Indeed, this thesis challenges the credibility of ESG measures on several occasions. For example, in Chapter 3, we show how there is a strong negative correlation between ES-scores and ESG controversy avoidance, which is not intuitive.

Future work should seek to answer the question of how best to measure true firm CSR. The definition of CSR and how best to capture it quantitatively have been the subject of great investigation. Nowadays, though, there is greater regulatory involvement around responsibility labelling. In the EU, a series of new regulation has been put forth to tackle greenwashing. For example, the Circular Economy Action Plan (CEAP), one of the European Green Deal's key building blocks, prohibits companies from claiming their products or services are "carbon neutral" without substantive proof (European Commission, 2022a). Since 2021, the EU's Sustainable Finance Disclosures Regulation (SFDR) has also been in force, requiring financial market participants to disclose, in good detail, how they are reducing their investments' negative impact on the environment and society before claiming 'sustainability' (European Commission, 2022b). Meanwhile, the Securities and Exchange Commission in the US has proposed rule changes that would make climate-related information "consistent, comparable, and decision-useful" for investors (SEC, 2022). In the near future, either established ESG ratings will reflect the information introduced through companies complying with these regulations, or new measures will emerge capturing only regulator-approved sustainability performance. Using more credible sustainability indicators is bound to add new value and perspective to research on the CSR-CFP link.

The research presented in this thesis would be of great value to investors, corporate managers, and regulators alike. We contribute to the work on the hypothesised resilience of socially responsible investments to market downturns, showing that this relationship is uncertain and inconsistent. Investors must look at ESG measures critically, as aggregate

ESG scores alone do not seem to capture true responsibility at every point in time. Investors may find a stronger link between CSR and financial returns when the most temporally relevant CSR dimension is looked at in isolation. We also warn corporations against window-dressing or greenwashing, not only because of new regulation that prohibits and prosecutes this, but also because our findings suggest that investors would be able to discern and penalise such behaviour. Finally, if socially responsible investors curb selling pressure during economic downturn, regulators wanting to stave off recessions may want to encourage the pursuit of genuine corporate social responsibility.

7 REFERENCES

Adam, A. M., & Shavit, T. (2008). How can a ratings-based method for assessing corporate social responsibility (CSR) provide an incentive to firms excluded from socially responsible investment indices to invest in CSR?. *Journal of Business Ethics*, *82*(4), 899-905.

Albuquerque, R., Koskinen, Y., Yang, S., & Zhang, C. (2020). Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. *The Review of Corporate Finance Studies*, *9*(3), 593-621.

Ali, W., Frynas, J. G., & Mahmood, Z. (2017). Determinants of corporate social responsibility (CSR) disclosure in developed and developing countries: A literature review. *Corporate Social Responsibility and Environmental Management*, *24*(4), 273-294.

Allen, T. D., Johnson, R. C., Kiburz, K. M., & Shockley, K. M. (2013). Work–family conflict and flexible work arrangements: Deconstructing flexibility. *Personnel Psychology*, *66*(2), 345-376.

Alon, T. M., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. (2020). The impact of COVID-19 on gender equality (No. w26947). *National Bureau of Economic Research*.

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, *74*(3), 87-103.

Anant, V., Banerjee, S., Boehm, J. & Li, K. (July 7, 2020). A dual cybersecurity mindset for the next normal. *McKinsey & Company*. https://www.mckinsey.com/business-functions/risk/our-insights/a-dual-cybersecurity-mindset-for-the-next-normal [Last Accessed: 23 April 2021].

Aouadi, A., & Marsat, S. (2018). Do ESG controversies matter for firm value? Evidence from international data. *Journal of Business Ethics*, *151*(4), 1027-1047.

Arouri, M., El Ghoul, S., & Gomes, M. (2021). Greenwashing and product market competition. *Finance Research Letters*, *42*, 101927.

Au, S., Dong, M. & Tremblay, A., (2019) Employee Flexibility, Exogenous Risk, and Firm Value. *Journal of Financial and Quantitative Analysis* (Forthcoming). *Available at SSRN 2883353*.Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir) responsible investments pay? New evidence from international ESG data. *The Quarterly Review of Economics and Finance*, *59*, 51-62.

Bae, K. H., El Ghoul, S., Gong, Z. J., & Guedhami, O. (2021). Does CSR matter in times of crisis? Evidence from the COVID-19 pandemic. *Journal of Corporate Finance*, *67*, 101876.

Bal, P. M., & Izak, M. (2020). Paradigms of flexibility: a systematic review of research on workplace flexibility. *European Management Review*, *18*(1), 37-50.

Baldauf, M., Garlappi, L., & Yannelis, C. (2020). Does climate change affect real estate prices? Only if you believe in it. *The Review of Financial Studies*, *33*(3), 1256-1295.

Balloch, A., Nicolae, A., & Philip, D. (2015). Stock market literacy, trust, and participation. *Review of Finance*, *19*(5), 1925-1963.

Barnett, M. L., & Salomon, R. M. (2006). Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, *27*(11), 1101-1122.

Barnett, M. L., & Salomon, R. M. (2012). Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal*, *33*(11), 1304-1320.

Barrero, J. M., Bloom, N., & Davis, S. J. (2020). Why Working From Home Will Stick. University of Chicago, Becker Friedman Institute for Economics Working Paper, (2020-174).

Bénabou, R., & Tirole, J. (2010). Individual and corporate social responsibility. *Economica*, *77*(305), 1-19.

Benson, K. L., & Humphrey, J. E. (2008). Socially responsible investment funds: Investor reaction to current and past returns. *Journal of Banking & Finance*, *32*(9), 1850-1859.

Berg, F., Kölbel, J. F., Pavlova, A., & Rigobon, R. (2021). ESG confusion and stock returns: Tackling the problem of noise. *Available at SSRN 3941514*.

Berg, F., Kölbel, J., & Rigobon, R. (2019) Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance* (Forthcoming). *Available at SSRN 3438533*.

Berkan, A., Leonardo, B., & Stefano, M. (2021). Media coverage, corporate social irresponsibility conduct, and financial analysts' performance. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1456– 1470.

Bernow, S., Klempner, B. & Magnin, C. (2017). From ‘why’ to ‘why not’: Sustainable investing as the new normal. *McKinsey & Company*. https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/from-why-to-why-not-sustainable-investing-as-the-new-normal [Last Accessed: 15 August 2022].

Bhatia, A., & Makkar, B. (2020). CSR disclosure in developing and developed countries: a comparative study. *Journal of Global Responsibility*, *11*(1), 1-26.

Bick, A., Blandin, A. & Mertens, K., Work from Home Before and after the Covid-19 Outbreak (2020). *CEPR* Discussion Paper No. DP15000. *Available at SSRN 3650114*.

Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2015). Does working from home work? Evidence from a Chinese experiment. *The Quarterly Journal of Economics*, *130*(1), 165-218.

Bloomberg (2021). Press Release: ESG Assets Rising to $50 Trillion Will Reshape $140.5 Trillion of Global AUM by 2025, Finds Bloomberg Intelligence. *Bloomberg*. https://www.bloomberg.com/company/press/esg-assets-rising-to-50-trillion-will-reshape-140-5-trillion-of-global-aum-by-2025-finds-bloomberg-intelligence.

Boffo, R., and R. Patalano (2020). ESG Investing: Practices, Progress and Challenges. OECD Paris. www.oecd.org/finance/ESG-Investing-Practices-Progress-and-Challenges.pdf [Last Accessed: 15 August 2022].

Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, *42*(3), 683-708.

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

Bourdeau-Brien, M., & Kryzanowski, L. (2017). The impact of natural disasters on the stock returns and volatilities of local firms. *The Quarterly Review of Economics and Finance, 63*, 259-270.

Breaugh, J. A., & Farabee, A. M. (2012). Telecommuting and flexible work hours: Alternative work arrangements that can improve the quality of work life. In *Work and Quality of Life* (pp. 251-274). Springer, Dordrecht.

Brynjolfsson, E., Horton, J. J., Ozimek, A., Rock, D., Sharma, G., & TuYe, H. Y. (2020). COVID-19 and remote work: An early look at US data (No. w27344). *National Bureau of Economic Research*.

Capelle-Blancard, G., & Petit, A. (2019). Every little helps? ESG news and stock market reaction. *Journal of Business Ethics*, *157*(2), 543-565.

Carbon Brief. (2022). https://www.carbonbrief.org/mapped-how-climate-change-affects-extreme-weather-around-the-world [Last Accessed: 18 April 2022]

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Carroll, A. B., & Shabana, K. M. (2010). The business case for corporate social responsibility: A review of concepts, research and practice. *International Journal of Management Reviews*, *12*(1), 85-105.

CCPI (2022). Climate Change Performance Index 2022. https://ccpi.org/download/climate-change-performance-index-2022-2.

CFTC (2020). Managing Climate Risk in the U.S. Financial System. Available at: https://tinyurl.com/4eaxbuaj [Last Accessed: April 23, 2021].

Chatjuthamard-Kitsabunnarat, P., Jiraporn, P., & Tong, S. (2014). Does religious piety inspire corporate social responsibility (CSR)? Evidence from historical religious identification. *Applied Economics Letters*, *21*(16), 1128-1133.

Chava, S. (2014). Environmental externalities and cost of capital. *Management Science*, *60*(9), 2223-2247.

Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic Management Journal*, *35*(1), 1-23.

Choi, D., Gao, Z., & Jiang, W. (2020). Attention to global warming. *The Review of Financial Studies*, *33*(3), 1112-1145.

Chow, R., Harootunian, H., & Miyoshi, M. (1996). Global/local: Cultural production and the transnational imaginary. *Duke University Press*.

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review*, *97*(1), 147-175.

CIPD (2019) Cross-Sector Insights on Enabling Flexible Working. Available at: https://www.cipd.co.uk/Images/flexible-working-guide-2019-v2_tcm18-58713.pdf.

Climate Bonds (2022). Explaining green bonds. Climate Bonds Initiative. https://www.climatebonds.net/market/explaining-green-bonds [Last Accessed: 12 May 2022].

Connors, S., Anderson-MacDonald, S., & Thomson, M. (2017). Overcoming the 'window dressing' effect: Mitigating the negative effects of inherent skepticism towards corporate social responsibility. *Journal of Business Ethics*, *145*(3), 599-621.

Coombs, J. E., & Gilley, K. M. (2005). Stakeholder management as a predictor of CEO compensation: Main effects and interactions with financial performance. *Strategic Management Journal*, *26*(9), 827-840.

Crane, A., McWilliams, A., Matten, D., Moon, J., & Siegel, D. (2008). *The Oxford University Handbook of CSR*.

Dahlsrud, A. (2008). How corporate social responsibility is defined: an analysis of 37 definitions. *Corporate Social Responsibility and Environmental Management*, *15*(1), 1-13.

Daugaard, D. (2020). Emerging new themes in environmental, social and governance investing: a systematic literature review. *Accounting & Finance*, *60*(2), 1501-1530.

Davis, K. (1973). The case for and against business assumption of social responsibilities. *Academy of Management Journal*, *16*(2), 312-322.

De Franco, C. (2020). ESG Controversies and Their Impact on Performance. *The Journal of Investing*, *29*(2), 33-45.

De Menezes, L. M., & Kelliher, C. (2011). Flexible working and performance: A systematic review of the evidence for a business case. *International Journal of Management Reviews*, *13*(4), 452-474.

Deegan, C. (2002). Introduction: The legitimising effect of social and environmental disclosures–a theoretical foundation. *Accounting, Auditing & Accountability Journal*, *15*(3), 282-311.

Demers, E., Hendrikse, J., Joos, P., & Lev, B. (2020) ESG Didn't Immunize Stocks During the COVID-19 Crisis, But Investments in Intangible Assets Did. *Journal of Business Finance & Accounting*, *48*(3-4), 433-462.

Devinney, T. (2009). Is the Socially Responsible Corporation a Myth? The Good, the Bad, and the Ugly of Corporate Social Responsibility. *Academy of Management Perspectives*, *23*(2), 44-56.

Dick, P. (2009). Bending over backwards? Using a pluralistic framework to explore the management of flexible working in the UK police service. *British Journal of Management*, *20*, S182-S193.

Diez-Cañamero, B., Bishara, T., Otegi-Olaso, J. R., Minguez, R., & Fernández, J. M. (2020). Measurement of corporate social responsibility: A review of corporate sustainability indexes, rankings and ratings. *Sustainability*, *12*(5), 2153.

Ding, W., Levine, R., Lin, C., & Xie, W. (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics*, *141*(2), 802-830.

Dingel, J. I., & Neiman, B. (2020). How many jobs can be done at home?. *Journal of Public Economics*, *189*, 104235.

Doktycz, C., & Abkowitz, M. (2019). Loss and damage estimation for extreme weather events: state of the practice. *Sustainability*, *11*(15), 4243.

Donadelli, M., Jüppner, M., Paradiso, A., & Ghisletti, M. (2020). Tornado activity, house prices, and stock returns. *The North American Journal of Economics and Finance*, *52*, 101162.

Du, S., Bhattacharya, C. B., & Sen, S. (2010). Maximizing business returns to corporate social responsibility (CSR): The role of CSR communication. *International Journal of Management Reviews*, *12*(1), 8-19.

Dutch News. (2021). Maas river keeps rising, death toll increases in Germany, Belgium. *Dutch News*. https://www.dutchnews.nl/news/2021/07/maas-river-keeps-rising-death-toll-increases-in-germany-belgium/ [Last Accessed: 18 April 2022].

Dyck, A., & Zingales, L. (2002). The corporate governance role of the media. (No. w9309). *National Bureau of Economic Research*.

Dyck, A., Morse, A., & Zingales, L. (2010). Who blows the whistle on corporate fraud?. *The Journal of Finance*, *65*(6), 2213-2253.

Eccles, R. G., and Klimenko, S. (2019). The Investor Revolution. Shareholders are getting serious about sustainability. *Harvard Business Review*, May-June 2019.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835-2857.

El Ghoul, S., Guedhami, O., Kwok, C. C., & Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital?. *Journal of Banking & Finance*, *35*(9), 2388-2406.

Elliot Gulliver-Needham, E. (2021). Schroders: Investors care more about the environment than ever. *Investment Week*. https://www.investmentweek.co.uk/news/4037925/schroders-investors-care-about-environment [Last Accessed: 12 May 2022].

Elliott, L., (2020) Number working from home in UK rises after government U-turn. *The Guardian*. https://www.theguardian.com/business/2020/oct/01/number-of-people-working-from-home-in-uk-rises-following-government-u-turn [Last Accessed: 10 December 2021].

Elsaid, A. M. (2012). The effects of cross-cultural work force diversity on employee performance in Egyptian pharmaceutical organizations. *Business and Management Research*, *1*(4), 162.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, *33*(3), 1184-1216.

Erhemjamts, O., Li, Q., & Venkateswaran, A. (2013). Corporate social responsibility and its impact on firms' investment policy, organizational structure, and performance. *Journal of Business Ethics*, *118*(2), 395-412.

European Commission (2001). Green Paper: Promoting a European framework for Corporate Social Responsibility https://ec.europa.eu/commission/presscorner/detail/en/DOC_01_9 [Last Accessed: 15 August 2022].

European Commission (2021). European Green Deal. https://ec.europa.eu/clima/eu-action/european-green-deal_en [Last Accessed: 18 April 2022].

European Commission (2022a). Circular economy action plan. https://ec.europa.eu/environment/strategy/circular-economy-action-plan_en [Last Accessed: 12 May 2022].

European Commission (2022b). Sustainability-related disclosure in the financial services sector. https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/sustainability-related-disclosure-financial-services-sector_en [Last Accessed: 12 May 2022].

Fahlenbrach, R., Rageth, K., & Stulz, R. M. (2020). How valuable is financial flexibility when revenue stops? Evidence from the COVID-19 crisis (No. w27106). *National Bureau of Economic Research*.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. The Journal of Finance, *47*(2), 427-465.

Fama, E. F., & French, K. R. (2021). The cross-section of expected stock returns (pp. 349-391). *University of Chicago Press*.

Favilukis, J. Y., Lin, X., Sharifkhani, A., & Zhao, X. (2020). Labor Force Telework Flexibility and Asset Prices: Evidence from the COVID-19 Pandemic. Georgetown McDonough School of Business Research Paper. Available at SSRN 3693239.

Fernandez-Feijoo, B., Romero, S., & Ruiz, S. (2014). Effect of stakeholders' pressure on transparency of sustainability reports within the GRI framework. *Journal of Business Ethics, 122*(1), 53-63.

Finder UK (2021). Working from home (WFH) statistics 2021. https://www.finder.com/uk/working-from-home-statistics [Last Accessed: April 23, 2021].

Folger-Laronde, Z., Pashang, S., Feor, L., & ElAlfy, A. (2022). ESG ratings and financial performance of exchange-traded funds during the COVID-19 pandemic. *Journal of Sustainable Finance & Investment*, *12*(2), 490-496.

Friedman, M. (1970). A Friedman doctrine: The social responsibility of business is to increase its profits. *The New York Times Magazine*, *13*(1970), 32-33.

Fukuyama, F. (1995). Trust: The social virtues and the creation of prosperity (Vol. 99). *New York: Free press*.

Galant, A., & Cadez, S. (2017). Corporate social responsibility and financial performance relationship: a review of measurement approaches. *Economic research-Ekonomska istraživanja*, *30*(1), 676-693.

Garel, A., & Petit-Romec, A. (2021). Investor rewards to environmental responsibility: Evidence from the COVID-19 crisis. *Journal of Corporate Finance*, *68*, 101948.

Geczy, C., Stambaugh, R. F., & Levin, D. (2005). Investing in socially responsible mutual funds. *Available at SSRN 416380*.

Georgarakos, D., & Pasini, G. (2011). Trust, sociability, and stock market participation. *Review of Finance*, *15*(4), 693-725.

Gertler, M., & Bernanke, B. (1989). Agency costs, net worth and business fluctuations. In *Business Cycle Theory*. Edward Elgar Publishing Ltd.

Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, *66*, 101889.

Godfrey, P. C., Merrill, C. B., & Hansen, J. M. (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic Management Journal*, *30*(4), 425-445.

Goyal, A., & Santa-Clara, P. (2003). Idiosyncratic risk matters!. *The Journal of Finance*, *58*(3), 975-1007.

Graham, J. R., Li, S., & Qiu, J. (2008). Corporate misreporting and bank loan contracting. *Journal of Financial Economics*, *89*(1), 44-61.

Griffin, J. J., & Mahon, J. F. (1997). The corporate social performance and corporate financial performance debate: Twenty-five years of incomparable research. *Business & Society*, *36*(1), 5-31.

Guiso, L. Sapienza, P., & Zingales, L. (2015). The value of corporate culture. *Journal of Financial Economics*, *117*(1), 60-76.

Guiso, L., Sapienza, P., & Zingales, L. (2004). The role of social capital in financial development. *American Economic Review*, *94*(3), 526-556.

Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes?. *Journal of Economic Perspectives*, *20*(2), 23-48.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *The Journal of Finance*, *63*(6), 2557-2600.

Guiso, L., Sapienza, P., & Zingales, L. (2011). Civic capital as the missing link. In *Handbook of Social Economics* (Vol. 1, pp. 417-480). North-Holland.

Hail, L., Shawn, K., & Zhang, R. X. (2021). How do managers greenwash? Evidence from earnings conference calls. Working Paper.

Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, *49*(6), 62-66.

Heathfield, S. (2021). The Pros and Cons of a Flexible Work Schedule. *The Balance Careers*. https://www.thebalancecareers.com/advantages-and-disadvantages-of-flexible-work-schedules-1917964#employee-flexibility [Last Accessed: April 23, 2021].

Henriques, I., & Sadorsky, P. (1999). The relationship between environmental commitment and managerial perceptions of stakeholder importance. *Academy of Management Journal*, *42*(1), 87-99.

Hill, E.J., Grzywacz, J. G., Allen, S., Blanchard, V. L., Matz-Costa, C., Shulkin, S., & Pitt-Catsouphes, M. (2008). Defining and conceptualizing workplace flexibility. *Community, Work and Family*, *11*(2), 149-163.

Hoepner, A. G., & Schopohl, L. (2018). On the price of morals in markets: An empirical study of the Swedish AP-Funds and the Norwegian Government Pension Fund. *Journal of Business Ethics*, *151*(3), 665-692.

Hoepner, A. G., & Schopohl, L. (2020). State pension funds and corporate social responsibility: do beneficiaries' political values influence funds' investment decisions?. *Journal of Business Ethics*, *165*(3), 489-516.

Hoepner, A. G., Oikonomou, I., Sautner, Z., Starks, L. T., & Zhou, X. (2018). ESG shareholder engagement and downside risk. *AFA 2018, European Corporate Governance Institute* – Finance Working Paper No. 671/2020. *Available at SSRN 2874252.*

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15-36.

Hong, H., & Kostovetsky, L. (2012). Red and blue investing: Values and finance. *Journal of Financial Economics*, *103*(1), 1-19.

Hong, H., Li, F. W., & Xu, J. (2019). Climate risks and market efficiency. *Journal of Econometrics*, *208*(1), 265-281.

Hsu, P. H., Li, K., & Tsou, C. Y. (2021). The pollution premium. Available at SSRN 3578215.

Hunt, V., Simpson, B., and Yamada, Y. (2022). The case for stakeholder capitalism. *McKinsey & Company*. https://www.value-match.co.uk/wp-content/uploads/McKinsey-and-Company-The-Case-for-Stakeholder-Capitalism.pdf [Last Accessed: 18 April 2022].

Il Sole 24 Ore. (2020). Change in value during coronavirus outbreak of selected stock market indices worldwide from January 1 to March 18, 2020 [Graph]. *Statista*. https://www.statista.com/statistics/1105021/coronavirus-outbreak-stock-market-change [Last Accessed: 12 May 2022].

Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon tail risk. *The Review of Financial Studies*, *34*(3), 1540-1571.

In, S. Y., & Schumacher, K. (2021). Carbonwashing: ESG Data Greenwashing in a Post-Paris World. In *Settling Climate Accounts* (pp. 39-58). Palgrave Macmillan, Cham.

ISO (2010). ISO 26000 Social Responsibility. https://www.iso.org/iso-26000-social-responsibility.html [Last Accessed: 15 August 2022].

ITUC (2021). 2021 ITUC Global Rights Index. International Trade Union Confederation. https://files.mutualcdn.com/ituc/files/ITUC_GlobalRightsIndex_2021_EN_Final.pdf [Last Accessed: 12 May 2022].

Jacobs, B. I., & Levy, K. N. (2022). The Challenge of Disparities in ESG Ratings. *The Journal of Impact and ESG Investing*, *2*(3), 107-111.

JBA Risk (2021). Summer Floods in Europe 2021. *JBA Risk: The Flood People*. https://www.jbarisk.com/flood-services/event-response/summer-floods-in-europe-2021/ [Last Accessed: 18 April 2022].

Jones, G. H., Jones, B. H., & Little, P. (2000). Reputation as reservoir: Buffering against loss in times of economic crisis. *Corporate Reputation Review*, *3*(1), 21-29.

Jost, S., Erben, S., Ottenstein, P., & Zülch, H. (2022). Does corporate social responsibility impact mergers & acquisition premia? New international evidence. *Finance Research Letters*, *46*, 102237.

Kahn, L., Lange, F., & Wiczer, D. (2020). Labor Demand in the time of COVID-19: Evidence from vacancy postings and UI claims. (No. w27061). *National Bureau of Economic Research*.

Kaiser, L., & Welters, J. (2019). Risk-mitigating effect of ESG on momentum portfolios. *The Journal of Risk Finance*, *20*(5), 542-555.

Kang, C., Germann, F., & Grewal, R. (2016). Washing Away Your Sins? Corporate Social Responsibility, Corporate Social Irresponsibility, and Firm Performance. *Journal of Marketing*, *80*(2), 59–79.

Kaye, A. D., Okeagu, C. N., Pham, A. D., Silva, R. A., Hurley, J. J., Arron, B. L. et al. (2020). Economic Impact of COVID-19 Pandemic on Health Care Facilities and Systems: International Perspectives. *Best Practice & Research Clinical Anaesthesiology*, *35*(3), 293-306.

Kenneth (K.) French Data Library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html [Last Accessed: 01 Jan 2022].

Kim, E. H., & Lyon, T. P. (2015). Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, *26*(3), 705-723.

Kim, S. (2019). The process model of corporate social responsibility (CSR) communication: CSR communication and its relationship with consumers' CSR knowledge, trust, and corporate reputation perception. *Journal of Business Ethics*, *154*(4), 1143-1159.

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, *112*(4), 1251-1288.

Kotchen, M., & Moon, J. J. (2012). Corporate social responsibility for irresponsibility. *The BE Journal of Economic Analysis & Policy*, *12*(1).

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, *115*(2), 304-329.

Kulak, F. A., & Tüzüner, V. L. (2020). A comparative analysis of flexible working patterns in Germany and Turkey. *International Journal of Research in Business and Social Science* (2147-4478), *9*(4), 01-14.

Lane, P. R. (2012). The European sovereign debt crisis. *Journal of Economic Perspectives, 26*(3), 49-68.

Lankoski, L. (2009). Differential economic impacts of corporate responsibility issues. *Business & Society, 48*(2), 206-224.

Leonardi, R., Nanetti, R. Y., & Putnam, R. D. (2001). Making democracy work: Civic traditions in modern Italy. Princeton, NJ: *Princeton University Press*.

Lexology (2021). The Evolution of ESG from CSR. https://www.lexology.com/library/detail.aspx?g=80bbe258-a1df-4d4c-88f0-6b7a2d2cbd6a [Last Accessed: 15 August 2022].

Li, J., Haider, Z. A., Jin, X., & Yuan, W. (2019). Corporate controversy, social responsibility and market performance: international evidence. *Journal of International Financial Markets, Institutions and Money*, *60*, 1-18.

Lindsey, L. A., Pruitt, S., & Schiller, C. (2021). The Cost of ESG Investing. *Available at SSRN 3975077*.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance, 72*(4), 1785-1824.

Lins, K. V., Volpin, P., & Wagner, H. F. (2013). Does family control matter? International evidence from the 2008–2009 financial crisis. *The Review of Financial Studies*, *26*(10), 2583-2619.

Liu, H., Yi, X., & Yin, L. (2021). The impact of operating flexibility on firms' performance during the COVID-19 outbreak: Evidence from China. *Finance Research Letters*, *38*, 101808.

Liu, O. and Mai, T., (2020) Employment during the COVID-19 Pandemic: Collapse and Early Recovery. *Available at SSRN 3682369*.

Lööf, H., & Stephan, A. (2019). The Impact of ESG on Stocks' Downside Risk and Risk Adjusted Return. No 477, Working Paper Series in Economics and Institutions of Innovation, *Royal Institute of Technology, Centre of Excellence for Science and Innovation Studies*.

Lopez-de-Silanes, F., La Porta, R. F., Shleifer, A., & Vishny, R. W. (1997). Trust in large Organizations. *The American Economic Review*, *87*(2), 333-338.

Mackey, A., Mackey, T. B., & Barney, J. B. (2007). Corporate social responsibility and firm performance: Investor preferences and corporate strategies. *Academy of Management Review*, *32*(3), 817-835.

Margolis, J., Elfenbein, H. & Walsh, J. (2009). Does it Pay to Be Good...And Does it Matter? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance. *Available at SSRN 1866371*.

Markos, S., & Sridevi, M. S. (2010). Employee engagement: The key to improving performance. *International Journal of Business and Management*, *5*(12), 89.

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, *27*(2), 483-504.

McCabe, R., Schmit, N., Christen, P. et al. (2020). Adapting hospital capacity to meet changing demands during the COVID-19 pandemic. *BMC Med*, *18*, 329.

McGrath, M. (2021a). US rejoins Paris accord: Biden's first act sets tone for ambitious approach. *BBC*. https://www.bbc.co.uk/news/science-environment-55732386 [Last Accessed: 18 April 2022].

McGrath, M. (2021b). Climate change: Extreme weather events are 'the new norm'. *BBC*. https://www.bbc.co.uk/news/science-environment-59105963 [Last Accessed: 18 April 2022].

McNall, L. A., Masuda, A. D., & Nicklin, J. M. (2009). Flexible work arrangements, job satisfaction, and turnover intentions: The mediating role of work-to-family enrichment. *The Journal of Psychology*, *144*(1), 61-81.

McWilliams, A., & Siegel, D. (2000). Corporate social responsibility and financial performance: correlation or misspecification?. *Strategic Management Journal*, *21*(5), 603-609.

Mecaj, A., & Bravo, M. I. G. (2014). CSR actions and financial distress: Do firms change their CSR behavior when signals of financial distress are identified?. *Modern Economy Scientific Research*, *5*, 259-271.

Meijer, M. M., & Schuyt, T. (2005). Corporate social performance as a bottom line for consumers. *Business & Society*, *44*(4), 442-461.

Miles, M. P., & Covin, J. G. (2000). Environmental marketing: A source of reputational, competitive, and financial advantage. *Journal of Business Ethics*, *23*(3), 299-311.

Minor, D., & Morgan, J. (2011). CSR as reputation insurance: Primum non nocere. *California Management Review*, *53*(3), 40-59.

Mongey, S., Pilossoph, L., & Weinberg, A. (2020). Which workers bear the burden of social distancing policies? (No. w27085). *National Bureau of Economic Research*.

Montenovo, L., Jiang, X., Rojas, F. L., Schmutte, I. M., Simon, K. I., Weinberg, B. A., & Wing, C. (2020). Determinants of disparities in COVID-19 job losses (No. w27132). *National Bureau of Economic Research*.

Moser, D. V., & Martin, P. R. (2012). A broader perspective on corporate social responsibility research in accounting. *The Accounting Review*, *87*(3), 797-806.

Murphy, P. J., & Coombes, S. M. (2009). A model of social entrepreneurial discovery. *Journal of Business Ethics*, *87*(3), 325-336.

Nollet, J., Filis, G., & Mitrokostas, E. (2016). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, *52*, 400-407.

Norman, G., Howard, A. M., Hardaway, W. K., and Reeves, A. B. (2021). ESG Rules: Will the UK Align With the EU? *Skadden*. https://www.skadden.com/insights/publications/2021/02/insights-special-edition-brexit/esg-rules-will-the-uk-align-with-the-eu [Last Accessed: 18 April 2022].

Oikonomou, I., Brooks, C., & Pavelin, S. (2012). The impact of corporate social performance on financial risk and utility: A longitudinal analysis. *Financial Management*, *41*(2), 483-515.

Oikonomou, I., Brooks, C., & Pavelin, S. (2014). The financial effects of uniform and mixed corporate social performance. *Journal of Management Studies*, *51*(6), 898-925.

Olson, E. (2013). BSR and the State of CSR: What We Mean When We Say 'CSR'. Business for Social Responsibility. https://www.bsr.org/en/our-insights/blog-view/bsr-and-the-state-of-csr-what-we-mean-when-we-say-csr [Last Accessed: 15 August 2022].

Origo, F., & Pagani, L. (2008). Workplace flexibility and job satisfaction: some evidence from Europe. *International Journal of Manpower*, *29*, 539-566.

Papanikolaou, D., & Schmidt, L. D. (2020). Working remotely and the supply-side impact of COVID-19 (No. w27330). *National Bureau of Economic Research*.

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021a). Dissecting green returns (No. w28940). *National Bureau of Economic Research*.

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021b). Sustainable investing in equilibrium. *Journal of Financial Economics*, *142*(2), 550-571.

Pearson, J. (2010). Are we doing the right thing? Leadership and prioritisation for public benefit. *Journal of Corporate Citizenship*, *37*, 37-40.

Peng, C. W., & Yang, M. L. (2014). The effect of corporate social performance on financial performance: The moderating effect of ownership concentration. *Journal of Business Ethics*, *123*(1), 171-182.

Qiang, C. Z., Saurav, A., and Viney, B. (2021). Global investors shift focus to sustainability amid push for a green recovery. *World Bank*. https://blogs.worldbank.org/psd/global-investors-shift-focus-sustainability-amid-push-green-recovery [Last Accessed: 18 April 2022].

Ramani, V. (2020). Climate Change Is a Systemic Financial Risk. *The Regulatory Review*. https://www.theregreview.org/2020/11/04/ramani-climate-change-systemic-financial-risk/ [Last Accessed: 18 April 2022].

Refinitiv (2021), Environmental, Social and Governance (ESG) Scores from Refinitiv, (February 2021). https://shorturl.at/wyMX9 [Last accessed: 28 Mar 2022].

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, *14*(3), 302-322.

Renneboog, L., Ter Horst, J., & Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, *20*(4), 562-588.

Reuters (2021). Asset size of sustainable funds worldwide as of September 2021, by region (in billion U.S. dollars) [Graph]. *Statista*. https://www.statista.com/statistics/1296334/sustainable-funds-asset-size-by-region.

Riedl, A., & Smeets, P. (2017). Why do investors hold socially responsible mutual funds?. *The Journal of Finance*, *72*(6), 2505-2550.

Rodgers, W., Choy, H. L., & Guiral, A. (2013). Do investors value a firm's commitment to social activities?. *Journal of Business Ethics*, *114*(4), 607-623.

Roxby, P. (2020). Why is it being called a pandemic now? *BBC*. https://www.bbc.co.uk/news/world-51839944 [Last Accessed: 12 May 2022].

Sacconi, L. & D. Antoni, G. (2011). Social Capital, Corporate Responsibility, Economic Behaviour and Performance. Palgrave MacMillan, New York, NY.

Sánchez, A. M., Pérez, M. P., de Luis Carnicer, P., & Jiménez, M. J. V. (2007). Teleworking and workplace flexibility: a study of impact on firm performance. *Personnel Review*, *36*, 42–64.

Schanzenbach, M. M., & Sitkoff, R. H. (2020). ESG investing: Theory, evidence, and fiduciary principles. *Journal of Financial Planning* (Forthcoming). *Available at SSRN 3684979*.

Schmidt, N., Pleitgen, F., Wojazer, B., and Ravindran, J. (2021). More than 150 people still missing in German floods unlikely to be found, officials fear. *CNN*. https://edition.cnn.com/2021/07/22/europe/germany-belgium-europe-floods-death-climate-intl/index.html [Last Accessed: 18 April 2022].

Schnietz, K. E., & Epstein, M. J. (2005). Exploring the financial value of a reputation for corporate social responsibility during a crisis. *Corporate Reputation Review*, *7*(4), 327-345.

Scrivens, K., & Smith, C. (2013). Four interpretations of social capital: an agenda for measurement. *OECD* Statistics Working Papers, 2013/06.

SEC (2022). Press Release: SEC Proposes Rules to Enhance and Standardize Climate-Related Disclosures for Investors. *U.S. Securities and Exchange Commission*. https://www.sec.gov/news/press-release/2022-46 [Last Accessed: 12 May 2022].

Servaes, H., & Tamayo, A. (2013). The impact of corporate social responsibility on firm value: The role of customer awareness. *Management Science*, *59*(5), 1045-1061.

Shan, C., & Tang, D. Y. (2020). The value of employee satisfaction in disastrous times: Evidence from COVID-19. *Available at SSRN 3560919*.

Sheehy, B. (2015). Defining CSR: Problems and solutions. *Journal of Business Ethics*, *131*(3), 625-648.

Shen, N., Au, K., & Li, W. (2019). Strategic alignment of intangible assets: The role of corporate social responsibility. *Asia Pacific Journal of Management*, 1-21.

Sherwood, M. W., & Pollard, J. L. (2018). The risk-adjusted return potential of integrating ESG strategies into emerging market equities. *Journal of Sustainable Finance & Investment*, *8*(1), 26-44.

Steffen, T. (2021). ESG scores: An outdated concept. *Responsible Investor*. https://www.responsible-investor.com/articles/esg-scores-an-outdated-concept [Last Accessed: April 23, 2021].

Sun, N., Salama, A., Hussainey, K., & Habbash, M. (2010). Corporate environmental disclosure, corporate governance and earnings management. *Managerial Auditing Journal*, *25*(7): 679– 700.

Surroca, J., Tribó, J. A., & Waddock, S. (2010). Corporate responsibility and financial performance: The role of intangible resources. *Strategic Management Journal*, *31*(5), 463-490.

Tang, D. Y., Yan, J., & Yao, C. Y. (2021). The determinants of ESG ratings: Rater ownership matters. Available at SSRN 3889395.

Turker, D. (2009). Measuring corporate social responsibility: A scale development study. *Journal of Business Ethics*, *85*(4), 411-427.

UK Health Security Agency (2021). COVID-19: epidemiology, virology and clinical features. *UK Government*. https://www.gov.uk/government/publications/wuhan-novel-coronavirus-background-information/wuhan-novel-coronavirus-epidemiology-virology-and-clinical-features [Last Accessed: 12 May 2022].

UN Industrial Development Organisation (2022). What is CSR? https://www.unido.org/our-focus/advancing-economic-competitiveness/competitive-trade-capacities-and-corporate-responsibility/corporate-social-responsibility-market-integration/what-csr [Last Accessed: 15 August 2022].

UNDRR (2020). Human cost of disaster. Centre for Research on the Epidemiology of Disasters and United Nations Office for Disaster Risk Reduction. https://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019.

United Nations (2022). Global Issues: Climate Change. https://www.un.org/en/global-issues/climate-change [Last Accessed: 12 May 2022].

UNPRI (2021). Signatory Update for October to December 2021. https://www.unpri.org/download?ac=14962 [Last Accessed: 18 April 2022].

Van Beurden, P., & Gössling, T. (2008). The worth of values–a literature review on the relation between corporate social and financial performance. *Journal of Business Ethics*, *82*(2), 407.

Van Duuren, E., Plantinga, A., & Scholtens, B. (2016). ESG integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics*, *138*(3), 525-533.

Venturini, A. (2022). Climate change, risk factors and stock returns: A review of the literature. *International Review of Financial Analysis*, *79*, 101934.

Verheyden, T., Eccles, R. G., & Feiner, A. (2016). ESG for all? The impact of ESG screening on return, risk, and diversification. *Journal of Applied Corporate Finance*, *28*(2), 47-55.

Vilanova, M., Lozano, J. M., & Arenas, D. (2009). Exploring the nature of the relationship between CSR and competitiveness. *Journal of Business Ethics*, *87*(1), 57-69.

Waddock, S. A., & Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic Management Journal*, *18*(4), 303-319.

Wang, L., & Kutan, A. M. (2013). The impact of natural disasters on stock markets: Evidence from Japan and the US. *Comparative Economic Studies*, *55*(4), 672-686.

Weber, M. (2008). The business case for corporate social responsibility: A company-level measurement approach for CSR. *European Management Journal*, *26*(4), 247-261.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838.

WHO (2020). Archived: WHO Timeline - COVID-19. *World Health Organization*. https://www.who.int/news/item/27-04-2020-who-timeline---covid-19 [Last Accessed: 12 May 2022].

Withisuphakorn, P., & Jiraporn, P. (2019). Deciphering the motives behind corporate social responsibility (CSR) using managerial ownership: evidence from heteroskedastic identification. *Applied Economics Letters*, *26*(12), 963-970.

Wood, D. J., & Jones, R. E. (1995). Stakeholder mismatching: A theoretical problem in empirical research on corporate social performance. *The International Journal of Organizational Analysis*, *3*, 229–267.

World Business Council for Sustainable Development (2000). Corporate Social Responsibility: Making Good Business Sense (Geneva, Switzerland).

Worldometer (2022). Number of novel coronavirus (COVID-19) deaths worldwide as of May 9, 2022, by country [Graph]. *Statista*. https://www.statista.com/statistics/1093256/novel-coronavirus-2019ncov-deaths-worldwide-by-country.

Worthington, A. C. (2008). The impact of natural events and disasters on the Australian stock market: A GARCH-M analysis of storms, floods, cyclones, earthquakes and bushfires. *Global Business and Economics Review*, *10*(1), 1-10.

Wright, P. D., & Bretthauer, K. M. (2010). Strategies for addressing the nursing shortage: Coordinated decision making and workforce flexibility. *Decision Sciences*, *41*(2), 373-401.

Yoon, B., Lee, J. H., & Byun, R. (2018). Does ESG performance enhance firm value? Evidence from Korea. *Sustainability*, *10*(10), 3635.

Yu, E. P. Y., Van Luu, B., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance*, *52*, 101192.

Yue, C. A., Tao, W., & Ferguson, M. A. (2022). The joint effect of corporate social irresponsibility and social responsibility on consumer outcomes. *European Management Journal (In press)*.

Zak, P. J., & Knack, S. (2001). Trust and growth. *The Economic Journal*, *111*(470), 295-321.