Christian Kreuzer

Financial aspects of sustainability: Drivers and consequences of corporate social responsibility and irresponsibility

University of Regensburg
Regensburg, Germany

Financial aspects of sustainability:
Drivers and consequences of corporate social responsibility and irresponsibility

Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg

vorgelegt von: Christian Kreuzer

Berichterstatter: Prof. Dr. Gregor Dorfleitner, Prof. Dr. Sven Bienert

Tag der Disputation: 21.07.2022

Christian Kreuzer

Financial aspects of sustainability: Drivers and consequences of corporate social responsibility and irresponsibility

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
Doktor der Wirtschaftswissenschaft (Dr. rer. pol.)

Advisors:
Prof. Dr. Gregor Dorfleitner
Prof. Dr. Sven Bienert

University of Regensburg
Regensburg, Germany

"One of the endlessly alluring aspects of mathematics is that its thorniest paradoxes have a way of blooming into beautiful theories."

Philip J. Davis, American academic applied mathematician, 1964.

Acknowledgements

First of all, I owe my gratitude to my advisor, Prof. Dr. Gregor Dorfleitner, for his continuous supervision, advice, and encouraging comments throughout my dissertation at the Department of Finance at the University of Regensburg. I am particularly grateful for his trust and for the opportunity with which he provided me to pursue my Ph.D. I am also thankful towards Prof. Dr. Sven Bienert for being my second advisor. Inspiring discussions with both and their beneficial suggestions profoundly improved my research.

I would also like to thank my colleagues at the Department of Finance for their helpful suggestions and our fruitful discussions.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The issue of sustainability is and continues to be one of the greatest challenges in our days. In essence, sustainability covers the union of economic, environmental, and social considerations (Elkington, 1997) as well as—following the United Nations Brundtland commission (Brundtland, 1987)—the developments necessary to meet the needs of present society and simultaneously to ensure this ability for future generations. In particular the Paris Climate Agreement in 2015 and also the European Green Deal in 2019 formulated groundbreaking goals of sustainability for the upcoming decades. In all its facets, this topic affects not only every individual's everyday life but it is also changing the business environment for companies in a profound way. Above all, sustainable economic development reveals that, especially in developed countries, both society and politics increasingly as well as rigorously demand companies act socially responsibly through direct or indirect incentives (Matten and Moon, 2008).

In this context, the term of corporate social responsibility (CSR) is commonly understood as business activities that focus on the improvement of social welfare regardless of profit maximization (Liang and Renneboog, 2017) as well as actions that appear to enhance social good beyond a company's interests and also beyond that which is required by law (McWilliams and Siegel, 2001). It is a concept that is attracting increasing worldwide attention and has been studied from a practical and academic point of view for many years. In this regard, researchers as well as investors focus on the role of business in society and examine a variety of issues, likewise, whether and to what extend companies and organizations benefit from engaging in CSR (Carroll and Shabana, 2010).

Reinforcing and highlighting socially responsible activities usually results in a higher level of (perceived) corporate social performance (CSP), which is a rather quantified measure of the virtually isomorphic concept of corporate social responsibility. Thus, progressive companies pursue to set themselves apart from their competitors through CSR engagements (Jamali and Mirshak, 2007). In addition to socially motivated aspects that argue for a

commitment in CSR, it is equally important from an academic and business perspective to examine the topic of CSR. Here, the relation of CSR and corporate financial performance (CFP) deserves particular attention.

Even though corporate socially responsible behavior is gaining increasing attention in today's societal perception, acting socially responsible entails, following Lin-Hi and Müller (2013), not only the idea of doing "good" but also comprises the responsibility for avoiding "bad". This implies for example illegal, unethical, as well as socially irresponsible behavior, which is captured by the notion of corporate social irresponsibility (CSI).

This dissertation mainly focuses on financial aspects of sustainability through examining various issues of CSR and CSI. Thereby, it focuses on both the investor's view and a corporate perspective.

The remaining introductory chapter is organized as follows: The ensuing subsection provides an overview about literature fields on the CSR and CFP relation, followed by a discussion of the issue of CSI. After this subsection, aspects of socially responsible investing in the context of socially responsible mutual funds are examined. The contribution subsection highlights the contribution of this dissertation to the academic literature. Afterwards, four relevant research papers are summarized in greater detail, regarding the respective research questions, the datasets, the applied methodologies, the empirical results; the connection between these works is highlighted, too.
Chapters 2 to 5 present the four research papers that comprise the dissertation, whilst the last chapter concludes by highlighting implications as well as addressing limitations and aspects of further research.

The relationship between CSR and CFP

In the field of financial research, a key subject of CSR is its link to corporate financial performance. Since the 1970s, and thus long before issues of social responsibility as well as sustainability attracted public attention, a plethora of academic studies have examined the relationship between corporate social performance and corporate financial performance. In general, academic literature identifies the following four major theories regarding the relationship between CSR/CSP and CFP (see, e.g., Ullmann, 1985; McGuire et al., 1988; Preston and O'bannon, 1997; McWilliams and Siegel, 2000; Orlitzky et al., 2003).

CSR activities after prosperous CFP One strand of literature posits that companies engage in CSR activities because of a priori good financial performance (McGuire et al., 1988; Waddock and Graves, 1997; Orlitzky et al., 2003; Kang et al., 2016). This relationship is referred to as the *slack resources theory*. This theory proposes a positive and direct

linkage of prior financial performance to subsequent CSP (Waddock and Graves, 1997; Orlitzky et al., 2003). In this regard, the availability of slack financial resources, which is provided by prior financial performance, introduces the opportunity and freedom for a company to invest in CSR activities (Waddock and Graves, 1997). In this context, CSR activities are perceived as voluntary and philanthropic expenses that are not essential to the success of the company and, therefore, may mainly depend on the availability of excess funds (McGuire et al., 1988; Kang et al., 2016).

CSR activities after facing poor CFP Besides this approach of focusing on the CSR–CFP relation, an alternative perspective cannot be ignored: the *managerial opportunism theory*. This entails that managers may use CSR as a strategic tool to distract from their opportunistic behavior and to receive support from their stakeholders (Prior et al., 2008; Choi et al., 2013; Devinney et al., 2013; Martínez-Ferrero et al., 2016; Kang et al., 2016). In doing so, they exploit a socially-friendly image to disguise former value-destroying practices, which have damaged financial performance as well as stakeholders' interests (Prior et al., 2008). Moreover, Prior et al. (2008) highlight that abrupt improvements in a company's CSR may provide a warning signal to value-destroying practices and thus could even reinforce further negative impacts on a company's financial performance.

Good CFP due to CSR Like the slack resources theory, the *good management theory* (e.g., McGuire et al., 1988; McGuire et al., 1990; Waddock and Graves, 1997) also emphasizes a positive link between CFP and CSR, albeit in a different temporal ordering. Here, prior CSR provides subsequent CFP (Kang et al., 2016). In summary, this theory suggests that the engagement in CSR activities is part of good management (Kang et al., 2016). In doing so, a company's management establishes a sound relationship with stakeholders, resulting in positive perception of the firm. This may improve CFP due to increased levels of sales or decreased stakeholder management costs (McGuire et al., 1990; Waddock and Graves, 1997; Hull and Rothenberg, 2008). Furthermore, the research stream focuses on the conception of doing well by doing good, arguing that the costs of superior CSR activities are lower than the resulting benefits (Hull and Rothenberg, 2008; Kang et al., 2016). Waddock and Graves (1997) find evidence that CSR may actually be an advantage in competition and does not lead to a competitive disadvantage.

Poor CFP due to CSR A considerable portion of academic literature focuses on a win–win approach of CSP and CFP, according to which financial and social sustainability issues can be achieved simultaneously. However, other authors emphasize trade-offs and conflicts and therefore link the CFP–CSP relation to *trade-off theory* (Aupperle et al., 1985; Preston and O'bannon, 1997; Hahn et al., 2010), which mainly reflects the view of

Friedman (1970). This theory points out that CSR activities may siphon off resources from a company (in particular: capital), which leads to a relative disadvantage compared to less socially responsible competitors. Thus, higher CSR efforts ceteris paribus lower a company's financial performance in comparison to competitors (Preston and O'bannon, 1997).

In order to bridge these widely divergent views, some authors address the differentiation of when and under what circumstances companies should engage in CSR (Barnett, 2007; Rowley and Berman, 2000). Moreover, Tang et al. (2012) address the issue of CSR engagement strategies to demonstrate how companies should engage in corporate social responsibility. The findings identify that a firm's financial performance suffers through inconsistent engagement in CSR. Moreover, it is more beneficial to focus on one or several closely related segments of CSR than to engage in all of them simultaneously.

Overall, following the exhaustive meta-analysis of Friede et al. (2015), most of the more than $2,000$ studies considered indicate a nonnegative CSP–CFP relation. Revelli and Viviani (2015) conclude that there is neither real cost nor benefit of socially responsible investing, and further that the performance level depends on the methodology applied by researchers. In this respect, the heterogeneity of the results may to a large portion be ascribed to the issue of how to quantify CSP.

To overcome this issue, researchers often rely on independent data providers, such as Refinitiv, Sustainalytics, or MSCI-KLD, who engage in publishing objective and transparent rating metrics regarding the three pillars of environmental (E), social (S), and governance (G). Due to their comprehensibility and comparability, these ESG scores are used to evaluate companies' securities and especially stocks. Despite the benefits of ESG scores, some researchers address the discordance of literature by highlighting limitations and biases in the measurement of CSP (Dorfleitner et al., 2015; Revelli and Viviani, 2015; Chatterji et al., 2016; Drempetic et al., 2019). In particular, Capelle-Blancard and Monjon (2012) as well as Revelli and Viviani (2015) ascribe the academic discordance mainly to the issue of data-driven results. Dorfleitner et al. (2015) and Chatterji et al. (2016) illustrate a lack of homogeneous ESG measurement concepts, which even occurs among the large international ESG-rating institutions. Drempetic et al. (2019) highlight shortcomings in the way ESG ratings measure corporate sustainability. To manage issues of CSP measurements, it may be useful to add further dimensions to the evaluation process. One of these dimensions may be to consider aspects of irresponsible behavior.

The issue of Corporate Social Irresponsibility

The notion of CSI does not necessarily equate to poor CSR efforts, which may be indicated by low ESG ratings. Thus, CSI may be understood as an additional and, to a certain extent, independent "fourth" ESG dimension that requires further consideration. As the concepts of corporate social responsibility and corporate social irresponsibility are not perfectly inverse (Arora and Dharwadkar, 2011; Aouadi and Marsat, 2018), a multitude of issues remains open, including the fundamental CSR–CSI relation. To examine this linkage, the issue of how to measure CSI arises. To quantify CSI, investors and academics rely on ESG controversies scores, which are also published by data providers, such as Refinitiv, and are based on negative news, in particular "scandals", of companies from various media sources. Note that in everyday language, the term 'controversy' comprises two legitimate opposite perspectives. The notion of a scandal, however, commonly refers a clearly deplorable behavior. Since, in the context of firms, every controversially discussed subject damages the reputation in one way or another, both terms are used interchangeably. In this regard, we define CSI as companies' actions that involve illegal, unethical, as well as socially irresponsible behavior, which is inspired by the Refinitiv controversies score methodology.

One possible intuitive approach to CSI may be to assume that companies that perform well in CSR, e.g., when they exhibit a good ESG reputation, also behave ethically correctly and, therefore, perform well in CSI or, more restrictively, exhibit no CSI so as to avoid reputation-harming controversies. Hence these companies place particular emphasis on being perceived as socially responsible and also strictly avoid reputation-damaging controversies. This seems especially important for companies whose business model is heavily dependent on a socially responsible image as well as good public perception, for example consumer brands. In this case, good CSR and CSI performance may ensure ongoing sales as well as good stakeholder relations. Lenz et al. (2017) find, when evaluating a sample based on the KLD database, that CSI occurs in the majority of companies that engage in CSR. Interestingly, when examining this finding through using Refinitiv data which comprises a comparable number of companies to the sample that is examined in the work of Lenz et al. (2017), the results are quite different. In fact, the yearly proportion of companies that exhibit at least one controversy varies decreasingly between approximately 10–25% (see Table 1.1). Yet, one fact cannot be ignored: when examining further descriptive statistics of the best and worst ESG quartile, which consist of the 25% best (worst) rated companies regarding the Refinitiv ESG score in each year, the percentage of firms with at least one controversy is many times higher in the best ESG quartile than in the worst ESG quartile.

Therefore, a considerable portion of firms with good CSR performance exhibits poor CSI performance through the involvement in one or more corporate controversies. This indi-

cates an opposed CSR–CSI relation, namely reaching high CSR and simultaneously poor CSI performance, which I call *Janus–phenomenon* in the spirit of the mythical Roman god with two faces. One possible explanation may be that a company values CSR, but operates in a business environment that is prone to corporate controversies (e.g., mining, commodities, oil) and is thusly exposed to a higher risk of being involved in a scandal. Another explanation may be that unethical or dubious actions appear to be an appropriate form of behavior in the countries in which a company operates (see, e.g., Siemens AG's bribery of the Argentine government or Glencore's bribery, corruption, and price manipulation scandal in Africa and Latin America—both firms exhibit high ESG ratings during the respective time periods). Thus, companies seem to adopt this behavior to successfully run their business in these countries.

Poor CSR performances could also be accompanied with good CSI performance, including companies that reach low ESG ratings but exhibit no corporate controversies. One reason behind this may be that a company operates in an environment where their executives as well as stakeholders see no need to establish a socially responsible reputation (one may think of a machine manufacturers for mining companies or companies operating in the tobacco or gambling industry) and, therefore, widely neglect CSR activities. Another reason may be that companies use financial resources not for CSR but for other, potentially more profitable, activities (e.g., growth). Accordingly, these firms also try to avoid scandals that may damage their reputation and possibly also their financial performance.

Finally, it is conceivable that companies perform poorly in both CSR and CSI. One explanation for this could be that executives do not pay heed to CSR and CSI, possibly because they operate in a scandal-prone business environment and, at the same time, see no need to spend resources on CSR. As a result, these firms perform poorly in both. However, considering Table 1.1, this case only applies to few companies.

Nevertheless, when considering the linkage between CSR and CSI the temporal and causal order is so far largely neglected. In this regard, academic literature that examines the link between corporate social responsibility and corporate social irresponsibility highlights the following two mechanisms.

Penance mechanism This mechanism describes that prior CSI of a company in time $t-x, x \geq 1$ causes engagement in CSR in time t (Kang et al., 2016). In line with Kotchen and Moon (2012), companies engage in CSR activities in order to offset former CSI. The authors consider the term CSI as "a set of actions that increases externalized costs and/or promotes distributional conflicts" (Kotchen and Moon, 2012, p.2), which is basically a reversal of Heal's view of CSR. Hence, CSI may pose a financial liability that firms strive to minimize through the engagement in CSR activities (Kotchen and Moon, 2012).

Table 1.1: Firms with controversies.

	Percentage of firms with at least one controversy		
Year	Full dataset	Best ESG quartile	Worst ESG quartile
2002	0.2441	0.3793	0.1250
2003	0.2351	0.3489	0.1277
2004	0.1557	0.2957	0.0497
2005	0.1702	0.3425	0.0696
2006	0.1948	0.3816	0.0705
2007	0.1791	0.3740	0.0634
2008	0.2021	0.4543	0.0512
2009	0.1860	0.4087	0.0517
2010	0.2135	0.4225	0.0646
2011	0.2160	0.4289	0.0800
2012	0.2000	0.3702	0.0919
2013	0.2071	0.4021	0.0829
2014	0.2023	0.3909	0.0795
2015	0.1004	0.2324	0.0250
2016	0.1175	0.2764	0.0266
2017	0.1038	0.2427	0.0190
2018	0.1057	0.2451	0.0318
2019	0.1201	0.2777	0.0371
2020	0.1422	0.2829	0.0627
2021	0.1225	0.2494	0.0408

This table presents the percentage of firms in the full dataset ($N = 84,189$ observations) that exhibit at least one controversy in the respective fiscal year, measured by the number of controversies regarding 23 controversy topics from Refinitiv.

Insurance mechanism Here, potential corporate socially irresponsible behavior in time t should cause a firm to engage in CSR in time $t - x, x \geq 1$ (Kang et al., 2016). In this case, CSR activities may offer insurance-like protection to temper negative judgments or sanctions of future CSI (Godfrey et al., 2009). Thus, CSR may serve as an intangible asset for a company in times of crisis to mitigate the impact of misconduct (Schnietz and Epstein, 2005) and establishes a metaphorical reservoir of goodwill among stakeholders that provides the respective company with idiosyncrasy credits to insure against potential CSI (Kang et al., 2016).

Besides a fundamental link between CSR and CSI, other authors consider the interaction of CSR and CSI (Lenz et al., 2017; Price and Sun, 2017; Tang et al., 2012). Lenz et al. (2017) find evidence that the positive firm value effect of CSR is weakened by the occurrence of CSI. Furthermore, Price and Sun (2017) demonstrate that the involvement in CSI exhibits a longer lasting impact than the engagement in CSR. They conclude that those firms doing little in terms of both CSR and CSI—i.e., low CSR efforts and few or none CSI—perform

better than firms that exhibit high levels of CSR and CSI. Additionally, Tang et al. (2012) pointed out that a firm's financial performance suffers from sporadic, possibly desultory CSR engagements. Consequently, there is little hope for "bad" companies of engaging appropriately in CSR as well as taking the advantage of properly-managed CSR.

However, academic work and numerous real-life examples reveal that socially irresponsible behavior could not only result in direct negative consequences for companies and stakeholders, such as significant losses in market value (Karpoff et al., 2005), but also may considerably harm a company's reputation (Grappi et al., 2013). In view of this, the question arises as to why, or more precisely under what circumstances, companies behave irresponsibly in the first place. Unfortunately, academic literature focusing on these key-questions is still very rare.

Socially responsible investing and mutual funds

Ethically motivated investors who evaluate aspects of CSR and CSI may incorporate socially responsible investing (SRI) strategies, for instance exclusively selecting stocks with a high ESG rating (positive screening) or strictly excluding the so-called "sin stocks" (e.g., alcohol, tobacco, or gambling industry) from their investment decisions (negative screening), into their stock-selection process. SRI strategies continue to receive a rapid rise in interest in recent years. According to the US Forum for Sustainable and Responsible Investments (US SIF, 2020), the AUM experienced a sharp increase from $12 trillion in the US market alone in 2018 to $17.1 trillion at the beginning of 2020. This corresponds to an increase of over 42%.

In light of good management and the trade-off hypothesis, the question arises whether socially responsible investing exhibits a performance benefit for investors. The existing academic literature offers controversial results. While some authors ascertain positive effects of SRI on portfolios' performance (Kempf and Osthoff, 2007; Edmans, 2011), others observe neither performance benefits (Statman and Glushkov, 2009) nor negative performance (Hong and Kacperczyk, 2009).

Apart from an active portfolio selection, passive investors, who refuse to make investment decisions themselves, may choose investments into mutual funds and thereby outsource the stock selection process to fund managers. In this regard, one field of academic literature analyzes the financial performance persistence of mutual funds (Hendricks et al., 1993; Carhart, 1997; Cuthbertson et al., 2008; Muñoz et al., 2013). In summary, hitherto no evidence for significant sustained overperformance of fund managers has been detected in these research studies. Moreover, Cuthbertson et al. (2008) point out that bad performance of mutual funds can mainly be attributed to bad skills of fund managers. In light of this

finding, prosperous or disastrous financial performance of a fund mainly depends on the skills of the respective fund manager, which highlights the importance of the fund selection process for investors.

However, passive investors who reach for sustainability may choose socially responsible (SR) mutual funds and, therefore, rely on fund managers, to fulfill their requirements for sustainability. To meet expectations and demands regarding social responsibility, these funds claim to incorporate SR criteria into their portfolio construction process. Consequently, another major strand of the SRI literature deals with the financial performance of SR mutual funds compared to conventional mutual funds. In doing so, the majority of authors reports that no performance differences are achieved here (Statman, 2000; Bello, 2005; Bauer et al., 2005; Cortez et al., 2009; Utz and Wimmer, 2014). Thus, this may be seen as a win-win situation for ethically motivated passive investors.

Particularly ethical investors may focus more on sustainable performance than on financial performance once they are convinced of the socially responsible nature of SR funds (Barreda-Tarrazona et al., 2011). Hence, the financial performance of SR mutual funds and their SR performance must be seen as two aspects of overall performance in their decision metric. Especially ethically motivated passive investors are confident that their investments will continue to meet their ethical standards. Therefore, it is an important, albeit so far a rather minor field of academic literature, to examine the SR persistence of socially responsible mutual funds. In this regard, Wimmer (2013) demonstrates that the overall social performance of SR mutual funds is persistent only for a short duration and declines after a few years, indicating the need for ethical investors to maintain their investments on a regular basis.

However, social performance is often captured by ESG ratings, which can be susceptible to aspects of "greenwashing". Thus, in particular for ethically motivated passive investors, it may also be important that the funds they invest in avoid companies that exhibit CSI (e.g., environmental scandals or human rights violations). Interestingly, academic literature addressing the CSI persistence of SR mutual funds is rare.

Contribution

The main objective of this thesis is to analyze various aspects of ethical and unethical corporate behavior, mainly with regard to the evaluation from an investor's perspective, as well as to explore influences on corporate controversies. In doing so, this dissertation contributes to the academic literature on corporate social responsibility and corporate social irresponsibility. In particular, it places significant emphasis on the growing academic field regarding corporate scandals.

One part of this work focuses on the relation between CSP and CFP and investigates various SRI strategies that allow investors to incorporate a company-based evaluation of ethical and unethical corporate behavior in their stock-selection process by using ESG and ESG controversies scores. Since ESG ratings are publicly available, advocates of the efficient market hypothesis (Fama, 1965, 1970) may argue that investors could not generate an outperformance of ESG-based stock selection strategy, while other authors reject perfect information-efficient markets (Grossman, 1976; Grossman and Stiglitz, 1980). Since ESG and ESG controversies ratings are published ex post, they can only be priced in with a delay, which hinders information efficiency on stock markets. Moreover, some unethical corporate behavior remains concealed for a certain duration (see Volkswagen emission scandal). Consequently, the controversies score is only able to capture corporate misconduct once this questionable behavior is publicly disclosed. Due to this, this dissertation also investigates aspects of informational efficient markets with respect to the presence and absence of ethical and unethical corporate behavior and analyzes to what extent investors may benefit from CSP-based investment strategies. By using the Refinitiv ESG controversies score, this work directly examines the mid-to-long-term effects of corporate controversies, which are identified as the new dimension of ESG, on the CFP within the context of portfolio selection. Thereby, it addresses the issue of how to measure and evaluate CSI from an investor's perspective. This work is the first that investigates ESG scandals in the context of portfolio-building.

Aside from various approaches of portfolio selection, this thesis aims to gain insights into CSR and CSI persistence in the context of socially responsible mutual funds. In particular, this dissertation addresses the task of how to measure the level of CSI of mutual funds as well as how to assess funds in terms of their CSI performance in a comparable way. Further, this work aims to compare aspects of CSR and CSI performance of SR mutual funds and to investigate if there is a link between fund management fee and CSR as well as CSI performance.

Apart from the fundamental CSR–CFP relation as well as the examination of SR persistence, a plethora of academic papers focus on potential nation-level drivers and company-specific motivations to engage in CSR (e.g., Ioannou and Serafeim, 2012; Liang and Renneboog, 2017; Baldini et al., 2018). The issue of potential drivers of corporate social irresponsibility remains largely unsolved in academic literature. Given this, another primary objective of this dissertation is to identify nation-level and company-related determinants which may influence the occurrence of corporate controversies. Furthermore, it sheds light onto the conceptual issue of what it takes to create a corporate controversy. Besides, questions of when a company adopts ethical behavior and on the role of society in disclosing unethical corporate activities are considered. Aside from societal, cultural, political, and firm-related variables this thesis aims to examine the role of board structure characteristics

in the context of a firm's irresponsible corporate behavior. Consequently, this work breaks new ground in CSI literature by explaining cross-national and intra-industrial differences between patterns of unethical corporate behavior. By developing a new theoretical model regarding the occurrence of corporate controversies, it establishes a link between unethical behavior and institutional as well as legitimacy theory.

This thesis is composed of four independent research papers with several co-authors.

Ch. 2 ESG controversies and controversial ESG: About silent saints and small sinners.
Publication status: Published in Journal of Asset Management

Ch. 3 How socially irresponsible are socially responsible mutual funds? A persistence analysis.
Publication status: Published in Finance Research Letters

Ch. 4 It's not only size that matters: On the influence of policy, society, culture, and firm characteristics on corporate controversies.
Publication status: Under review (resubmitted)

Ch. 5 Board responsibility for irresponsibility: The influence of board structure on corporate scandals.
Publication status: Major revision

Examining the relationship between CSP and CFP in the context of portfolio selection

The first article is a contribution to the discourse in academic literature on the relationship between corporate social performance and corporate financial performance. This relationship has been widely examined and discussed in academic research (e.g., Revelli and Viviani, 2015; Friede et al., 2015). From a stock-market-based perspective, some authors also investigate returns of ESG portfolios (Kempf and Osthoff, 2007; Halbritter and Dorfleitner, 2015). However, none of them analyze ESG scandals in their portfolio-building process. To the best of our knowledge, this research work is the first to investigate the impact of an ESG-based scandal within the scope of stock portfolio selection.

We examine positive screened portfolios, based on various cutoffs and different weighting strategies in the European, US, and global markets. The analysis in this article includes three distinct ratings: one that measures CSR, one for evaluating CSI, as well as a combined rating. These ratings allow for a straightforward implementation of a best-in-class approach, that does not a priori exclude any industries, such as "sin stocks". All three scores are derived from Refinitiv (formerly known as Thomson Reuters): first, to evaluate

a company's level of CSR, the *Refinitiv ESG score*, second, as a measurement of ESG-based controversies during a fiscal year, the *Refinitiv controversies score* and third, the *Refinitiv combined score*, which aggregates aspects of both ratings, the ESG controversies and the ESG score.

Our global dataset includes an average of approximately 2,500 companies per month in the years between 2002 and 2018. Through sorting stocks separately according to the three scores and selecting the best-rated and worst-rated stocks from each ranking, we construct several portfolios. In doing so, we consider a best-only and worst-only investment strategy as well as a best-minus-worst strategy, which entails the idea of investing long in the best-rated firms and simultaneously short in the worst-rated ones. To determine the risk-adjusted performance of the respective stock portfolios, our analysis is based on the Fama and French (2015) five-factor model.

Our results identify that, especially for small companies, the absence of controversies may be overlooked and/or incorrectly assessed in the share price. Thus, in particular from an investor's perspective, small firms which exhibit a "clean coat" with regard to controversies seem to be particularly profitable since these companies "fly under the radar". We additionally ascertain that equally-weighted portfolio strategies based on low ESG scores and combined scores also indicate significant outperformance, which may be ascribed to indications in support of the trade-off theory. A further approach regarding the issue of portfolio formation indicates that a so-called rank-weighted strategy, i.e., portfolio weightings are allocated by rank in ascending (or descending) order, exhibits new potential for investors, as implementing this strategy improves both alpha and level of significance within most of the portfolios under consideration.

We believe that this research paper offers a crucial contribution to academic literature regarding the relation between corporate financial performance and corporate social responsibility as well as irresponsibility. It is to be emphasized that the implementation of scoring methodologies to measure firm CSR as well as CSI in a portfolio-selection process is easy to replicate and may be of special interest for managers and investors.

Measuring the persistence of social responsibility and social irresponsibility of SR mutual funds

The aim of the second article is to analyze the persistence of CSI, measured by a controversies score, and CSP, measured by an environmental, social, governance (ESG) score, in socially responsible US mutual funds. It contributes to the academic literature regarding social performance in the context of mutual funds and also provides an assessment of CSR and CSI for investors. Moreover, this article offers another opportunity for in-

vestors to include CSI and CSR dimensions into investment decisions, besides an active portfolio-selection approach.

Although social performance of mutual funds is an important issue, most academic literature on the subject of mutual funds focuses on financial performance persistence (Hendricks et al., 1993; Carhart, 1997; Muñoz et al., 2013) or the comparison of financial performance of SR mutual funds and conventional mutual funds (e.g., Statman, 2000; Bauer et al., 2005; Cortez et al., 2009; Utz and Wimmer, 2014). While the overall sustainable performance of SR mutual funds was already discussed by Wimmer (2013), no academic literature has addressed the issue of examining CSI of SR equity funds so far. Hence, this work is the first to analyze and evaluate the persistence of US mutual equity funds regarding unethical corporate behavior as measured by an ESG controversies score based on an evaluation of the respective companies in their historical stock holdings.

In this article, we investigate the persistence of controversies scores and ESG scores in socially responsible US mutual funds by using a dataset that entails over 60 funds and over 400 fund compositions in the timeframe between 2003 and 2018. To compute the total ESG and controversies score of a particular fund's composition, we use the *Refinitiv ESG score* and *Refinitiv ESG controversies score* for the individual firms in the funds' holdings and calculate a weighted sum by using the latest available scores before the reporting date and the individual securities weights of the respective fund composition. In doing so, we follow Dorfleitner et al. (2012) when calculating the social return of portfolio holdings.

The procedures applied in this work are similar to the approach of Wimmer (2013). In this regard, we initially construct four equally weighted fund portfolios for each score, one for each quartile of the overall funds' scores, to categorize the funds concerning their ESG and controversies ratings on an annual basis. We examine the short-term persistence (one-year) as well as the mid- and long-term persistence (initial year and the subsequent four years). We further investigate whether high-paid SR fund managers achieve better ESG or controversies performance in comparison to lower-paid managers. Finally, analyze whether SR mutual funds which achieve high ESG ratings simultaneously perform well regarding the controversies scores and vice versa.

This article produces some noteworthy results. First, SR funds pursue clear and relatively consistent investment policies in the short, medium, and long term. Second, high-paid fund managers exhibit a better controversies-based social performance of their funds compared to lower-paid managers. However, when considering their funds' ESG ratings, high-paid managers are surpassed by lower-paid ones. Third, it is difficult for SR mutual funds to become leaders in both ESG and controversies ratings. Funds with high ESG ratings tend to achieve relatively low controversies ratings and vice versa. All in all, this work provides a proper first step toward a new aspect of ESG assessment of equity funds and is also of special interest for fund managers and ethically-motivated investors. This article also

illustrates differences between CSR and CSI. For a better understanding of CSI, especially from an investor's perspective, a closer examination of drivers of corporate controversies is addressed in the following paper (see chapter 4).

What influences corporate scandals

The aim of this article is to extend the understanding of irresponsible corporate behavior by examining drivers of CSR-based corporate controversies. In doing so, we investigate a variety of nation-level variables as well as company-specific determinants that may influence patterns of unethical corporate behavior, plus their disclosure. For a theoretical approach to a controversy, we introduce and discuss a conceptual model, which basically follows Thompson's definition of a media scandal (Thompson, 2005). The occurrence of corporate controversies comprises a two-stage process: first, the unethical corporate behavior of a firm, and second, the process of societal disclosure. This includes the perception of this corporate behavior, followed by indignation, which ultimately causes the need for action (Aouadi and Marsat, 2018; Weick et al., 2005).

This paper is the first to investigate an extensive spectrum of drivers of corporate controversies empirically. The analysis is based on a broad international data sample comprising over 5,700 companies located in 44 countries in the period of 2002–2017. It includes two metrics to measure a firm's corporate social irresponsibility: the Refinitiv ESG controversies score as well as the absolute number of controversies, which is a raw count variable. The dataset contains further country- and company-level variables that cover aspects of political, societal, cultural, and firm-related characteristics.

Our analyses are based on a within-between hybrid regression model (Allison, 2009; Schunck, 2013), which enables us to capture within and between effects while simultaneously keeping time-invariant effects, as well as a Tobit regression model.

The results indicate that firms will adopt ethical behavior if they operate in a business environment in which they are exposed to high levels of institutional pressure (institutional theory) or in which the involvement in a corporate scandal poses a direct threat to organizational legitimacy (legitimacy theory). Furthermore, our findings illustrate that firms operating in countries with both efficient enactment of laws and law enforcement as well as with low levels of corruption, higher levels of power distance, and in uncertainty-avoiding societies, tend to be less likely involved in a scandal. In countries with an individualistic culture, companies are more likely to become involved in a corporate controversy. Considering company-specific characteristics, companies that exhibit a good CSR reputation, larger firms, and high-attention firms increase the susceptibility to corporate controversies. In contrast, firms with a high level of dependency on capital markets tend to be less likely

involved in corporate controversies.

We believe that our work is a major contribution to CSR and CSI literature and lays the foundation for deeper insights into the topic of corporate scandals. Additionally, it may be of special interest to investors, managers, executives, and policy makers.

The influence of board structure on corporate scandals

The last paper investigates further firm-level drivers of CSI and focuses on the influence of board structure characteristics on the occurrence of corporate controversies. As many controversies stem from ethically or morally questionable decisions by executives, it seems a logical further step to examine the impact of a company's board, which represents the entity supervising a firm's executives, on corporate controversies from an academic perspective. Although the influence of board composition on corporate social irresponsibility is an important topic, only a few academic papers exist which address this relation (e.g., Jain and Zaman, 2020; Dharwadkar et al., 2021).

Our international dataset combines various data sources and comprises information from over 6,100 firms located in 44 countries in the time frame between 2002–2018. It includes political, societal, cultural, and firm-related determinants, a measure for CSI, which is inspired by the framework applied in Dorfleitner et al. (2021b), as well as various board variables. Again, to analyze our data, we use a hybrid regression model (Allison, 2009; Schunck, 2013).

Besides board characteristics that tend to increase a company's irresponsible corporate behavior, our findings also identify board variables that tend to lower such corporate behavior. In greater detail, firms with high levels of board skills and larger boards tend to be less likely involved in a corporate scandal, while companies associated with high board-level CSR efforts as well as high levels of board member affiliations tend to be more susceptible to the involvement in a corporate scandal. The results regarding the gender structure diversification of board members allow no clear conclusion.

This research work sheds further light on the understanding of the occurrence of corporate controversies and implements new approaches and ideas to extend the existing literature of corporate social irresponsibility. In addition to that, it highlights the importance of board skills and larger boards, which tend to prevent companies from being involved in a corporate controversy.

Chapter 2

ESG controversies and controversial ESG: About silent saints and small sinners

This research project is joint work with Gregor Dorfleitner (University of Regensburg) and Christian Sparrer (University of Regensburg). The paper has been published as:
Dorfleitner, G., Kreuzer, C., Sparrer, C. (2020), ESG controversies and controversial ESG: about silent saints and small sinners. *Journal of Asset Management*, 21 (5), 393–412.

Abstract Based on an extensive international dataset containing Thomson Reuters environmental, social and corporate governance (ESG) rating, as well as Thomson Reuters newest controversies and combined score of an average of 2,500 companies in the years 2002–2018, this article contributes to the existing discourse of the relationship between corporate social performance and corporate financial performance (CFP) by examining the Fama and French (2015) five-factor risk-adjusted performance of positive screened best and worst portfolios, based on a 10% cutoff, respectively for equally-, value- and rank-weighted strategies in the European, US, and global market. Furthermore, the controversies score allows us to examine the mid-to-long-term effects of scandals on the CFP without having to rely on the event study methodology. Even though a value-weighted strategy does not show any significant abnormal returns, we examined a significant outperformance for equally-weighted worst ESG-portfolios and best controversies strategies. These results strongly indicate that this is, on the one hand, driven by low-rated smaller companies ("small sinners") and clean-coated firms with regard to controversies ("silent saints") on the other hand. The findings hold for several robustness checks such as adjusting the cutoff rates or splitting the dataset across time.

Keywords ESG, corporate social responsibility, corporate social performance, controversy

2.1 Introduction

The interaction between corporate social performance (CSP) measured by ESG scores (which evaluate the performance of companies in their environmental, social or corporate governance pillars) and their corporate financial performance (CFP) has been the subject of academic research for many years with various findings. This paper is the first to examine the mid-to-long-term effects of controversies, as the new dimension of ESG, on the CFP of listed companies in a portfolio context. Furthermore, it determines the impact of different weighting strategies for high- and low-rated ESG and controversy portfolios.

Since the 1970s the matter of the relationship between CSP and CFP has been investigated by a pile of academic research. Revelli and Viviani (2015) report in their recent meta-analysis that the consideration of CSP in a portfolio leads to neither an under- nor an outperformance when compared with non-ESG-based investment strategies. Friede et al. (2015) conclude from their meta-analysis that approximately 90% of the more than 2,000 considered studies report a nonnegative relationship between CSP and CFP. This heterogeneity of the results can generally be ascribed to three issues, namely the question of how to measure CSP, the methods of stock selection and the question of how to define and measure CFP.

Addressing the first concern, some companies like Sustainalytics, MSCI-KLD or Asset4 specialize in issuing an ESG-based rating system and represent therefore as external and independent rating providers a transparent and reliable source of objective corporate social responsibility (CSR) measurements. Nevertheless, Capelle-Blancard and Monjon (2012) as well as Revelli and Viviani (2015) argue that the academic discordance can mainly be ascribed to the factor of data-driven results. Furthermore, Dorfleitner et al. (2015) and Chatterji et al. (2016) report a lack of homogeneous ESG measurement concepts, even among the large international ESG rating institutions.

To address the CSP measurement issue, our analysis includes three distinct ratings that represent industry-based percentile-ranked scores, which enable a simple implementation of a best-in-class approach and therefore do not discriminate any industry groups. The first one, the *Thomson Reuters ESG score* (in the following referred to as TR score), evaluates the CSR in various pillars, the *Thomson Reuters Controversies score* (in the following referred to as controversies score) measures the amount of ESG-based controversies a company encounters during a fiscal year, and finally, the *Thomson Reuters Combined score* (in the following referred to as combined score) aggregates ESG-related controversies and the TR score of a company.

Despite the fact that the controversies score finds its application within other financial research (see, for example, Park, 2018 and Vasilescu and Wisniewski, 2019), we still con-

tribute to the literature as we are the first ones to consider the extreme event of an ESG-based scandal within the context of portfolio selection.

The heterogeneity of academic results is strengthened even further by the use of various stock selection criteria. The most common and easy way in which an investor can implement a socially responsible investment (SRI) strategy is represented by socially responsible (SR) mutual funds. These funds claim to construct a portfolio based on SR selection criteria, such as selecting stocks with a high ESG rating (positive screening) or excluding the so-called "sin stocks" (tobacco, alcohol, arms or gambling industry) from their investment decisions (negative screening). The majority of the literature devoted to these type of investment strategies reports on no financial performance differences between SR and conventional mutual funds (see, i.e., Statman, 2000; Bauer et al., 2005; Bello, 2005; Kreander et al., 2005; Cortez et al., 2009; Utz and Wimmer, 2014). However, socially or ethically motivated value-driven investors in particular have to pay close attention to the shifting level of social responsibility of these SR funds. Wimmer (2013) finds that these funds are optimized towards their financial rather than their social performance and therefore the overall level of social performance of an SR fund is only persistent in the short run. Utz and Wimmer (2014) argue that, viewed from an individual stock level, neither SR mutual funds nor conventional funds differ greatly in terms of portfolio composition. This leads to the conclusion that SR mutual funds do not sustainably satisfy the needs of value-driven investors.

To overcome the stock selection problem, our analysis does not include SR funds, but rather selects stocks based on an ESG-ranking, allowing us to measure the CSR of a firm directly and therefore constructs long-term ESG-persistent portfolios by implementing a monthly rebalanced positive screening process following the ESG-based portfolio formation method of Kempf and Osthoff (2007). We construct a *best* and *worst* portfolio based on 10% cutoffs for ESG and controversy out- and underperformer in the sample, respectively. Additionally, the *best-minus-worst* zero-cost-investment strategy simply buys the outperformers and short sells the underperformers. Besides testing for the standard approach of value-weighted portfolios, we also conduct equally-weighted ones to better control for disparities between large and small firms. Furthermore, we implement a ranked weighting, which, given an ESG-based stock selection, allocates a higher weight to the respective stock the more extreme its score becomes.

Regarding the definition and measurement of CFP, researchers tend to use methods of two different directions. Whereas the first group, which represents an accounting-based view, defines CFP as the shift in earnings per share (EPS), operating profitability (return on equity (ROE), return on assets (ROA) or return on sales (ROS)) or net income, the second employs a stock-market-oriented perspective by applying (risk-adjusted) performance measurements such as abnormal returns, Sharpe Ratio or Tobin's Q. A common method

in the accounting-based direction comprises the implementation of a particular type of regression analysis. Qiu et al. (2016), for instance, regress the ROS of companies on their respective ESG score. Mervelskemper and Streit (2017) follow the valuation approach of Ohlson (1995) and add an ESG dimension to the model resulting in a regression of the market-to-book value of equity ratio on an ESG score. Van der Laan et al. (2008) implement a firm-fixed-effects regression to measure the influence of different CSP rating dimensions on the ROA and the EPS. In the stock-market-based perspective, factor models represent a common way in which to measure CFP as they have evolved from simple single-index models (like the CAPM) into a more appropriate approach like the Fama and French (2015) five-factor model. Kempf and Osthoff (2007) and Halbritter and Dorfleitner (2015), for example, align themselves in this group by implementing a Carhart (1997) four-factor model to estimate the abnormal returns of ESG portfolios. With a Fama and MacBeth (1973) regression, Halbritter and Dorfleitner (2015) also incorporate a cross-sectional approach as they regress the excess return of a certain company on its ESG score. Pintekova and Kukacka (2019) analyze the share prices of companies based on the Thomson Reuters combined score using a within-group fixed-effects model. Aouadi and Marsat (2018) utilize a fixed-effects model with dummy variables to estimate the relationship between Tobins' Q and an ESG score. Other studies, such as Auer (2016) and Auer and Schuhmacher (2016) who implement a Sharpe Ratio approach, rely on financial ratios. Event studies represent another noteworthy methodology, which is especially useful when analyzing the short-term impact of certain events (for example, the eventuation of a scandal). Among others, Lundgren and Olsson (2009) examine the effects of environmental-based scandals on firm value by applying a t test to the cumulative standardized abnormal return, whereas Krüger (2015) utilizes the cumulative abnormal return to show the impact of positive and negative ESG-related news separately on firm value. As these examples show, there is a wide variety of different methods and models for different purposes. A more stock-market-oriented perspective is especially suitable for an analysis from an investor's perspective as these methods better reflect the investors' perception of the impact of CSR on the future value of the company (see, i.e., Hillman and Keim, 2001; Gentry and Shen, 2010; Pintekova and Kukacka, 2019). Therefore, we align with the stock-market-oriented perspective and use the Fama and French (2015) five-factor model to calculate the risk-adjusted abnormal return. Furthermore, the use of the controversy score allows us to directly measure the mid-to-long-term effects of controversies on CFP without having to rely on the event study methodology.

Besides the academic disjointedness, SRI strategies have received a rapid rise in interest over the recent years. The global AUM, according to the Global Sustainable Investment Review GSIA (2018), grew significantly from \$22.89 trillion in 2016 to \$30.68 trillion in 2018, whereas, as reported by the U.S. Forum for Sustainable and Responsible Investments US SIF (2018), the AUM experienced a sharp increase from \$8.7 trillion in 2016 to \$12.0

trillion at the beginning of 2018 in the US market alone, which shows an almost 40% growth over two years. Furthermore, as mentioned by Crilly et al. (2012), the increasing pressure provided by various stakeholder groups forces companies to invest financial resources in CSR. Moreover, many investors pay close attention to the CSR or CSP of firms, whether they be value-driven investors trying to satisfy their altruistic needs or attempting to achieve abnormal returns by investing in firms with high ESG ratings.

Interestingly, within our results, we find a significant outperformance of up to almost 9% p.a. for the *worst* TR score portfolios for equally-weighted strategies as well as 7% p.a. for the equally-weighted *best* controversies score portfolios. These results show that investors should focus on low-rated smaller companies ("small sinners") and clean-coated firms with regard to controversies ("silent saints"). The implementation of a rank-weighted strategy instead of an equally-weighted one shows an improvement in alpha across nearly all tested strategies. Regarding the value-weighted strategies, no significant out- or underperformance can be found. These findings apply for different markets and hold true for various robustness checks.

This paper is organized as follows. Section 2.2 provides a short overview of the recent state of literature, while the data and methodology are discussed in Section 2.3. Section 2.4 presents our results. Section 2.5 implements several robustness checks, and Section 2.6 concludes.

2.2 Literature overview

This section provides an overview of the three perspectives regarding the relationship between CSP and CFP.

The first one indicates a positive relationship between the ESG score of a company and their respective CFP (see, i.e., Kempf and Osthoff, 2007; Statman and Glushkov, 2009; Auer, 2016; Pintekova and Kukacka, 2019) and is often referred to as *doing good while doing well*. This hypothesis holds true if the costs of socially responsible activities are overestimated or the respective benefits exceed the expectations of the managers and investors. This can be explained through the managerial myopia theory (see, i.e., Narayanan, 1985; Stein, 1988), where, on the one hand, managers tend to prefer decisions with a short-term profit rather than those that maximize long-term shareholder value, and short-term focused investors, on the other hand, who undervalue long-term benefits. Since the costs of socially responsible activities occur immediately, the benefits of those arise in the future. Therefore, the corresponding benefits are harder to predict and less attractive to short-term focused investors. Among others, Derwall et al. (2005) and Edmans (2011), who link the *doing good while doing well*-hypothesis with the managerial myopia theory, conclude

that short-term investors are unable (or unwilling) to price the long-term benefits of those activities correctly and therefore undervalue stocks of companies with high levels of engagement in environmental or social aspects, leading to higher returns in the long-run for the respective stocks when compared with other stocks. This idea of benefit manifestation in the long run is consistent with the findings of Dorfleitner et al. (2018), who conclude that the benefits of socially responsible activities (measured by the abnormal stock returns) are produced by unexpected additional cash flows which occur mid-to-long term. Pintekova and Kukacka (2019) divide the term of ESG-based activities into a primary and a secondary sector, whereas the first category refers to socially responsible activities which are closely related to the core business of the respective company. They can corroborate within their results, the point of view of *doing good while doing well* if the ESG-based activity is located in the primary sector.

The second approach reverts the above-mentioned relationship, which produces a view of *doing good but not well* (see, i.e., Boyle et al., 1997; Barnea and Rubin, 2010; Renneboog et al., 2008; Hong and Kacperczyk, 2009). This hypothesis holds true for many reasons. First of all, based on the idea of Barnea and Rubin (2010), socially responsible activities that represent lavish expenditures of managers motivated by personal benefits, such as public appreciation rather than the altruistic motive of non-financial utility, lead to a significant decrease in shareholder value and inferior financial performance. Thus, an agency problem occurs. As described by Krüger (2015), investors will react negatively (positively) to the announcement of socially responsible activities of firms with a high (low) amount of liquidity and can therefore be seen as wasteful investments. Furthermore, as stated by Heinkel et al. (2001), and Hong and Kacperczyk (2009), socially responsible investors and institutions which are subjected to social norm pressures (such as pension funds, universities, and religious organizations) exclude "sin stocks" from their investment decisions resulting in a lower demand, respectively, price and therefore a higher return in comparison with stocks which have a high ESG rating. Another reason supporting the *doing good but not well*-hypothesis is the trade-off theory stated by Aupperle et al. (1985). In the case of socially responsible investments, the theory argues that ESG-based activities exhaust financial resources which are lacking in other places. Thus, companies with a low level of expenditure on CSR achieve a competitive advantage in the long run, which may be especially relevant for smaller firms who are on a tighter budget. For small companies, the trade-off theory is strengthened even further by the findings of Aouadi and Marsat (2018). Since they examine the connection between firm visibility, CSP and CFP they conclude that only for high-attention firms (firms that are larger, more present in the media and more greatly observed by analysts), the ESG rating plays a role. In conclusion, if smaller firms invest in CSR, this could be seen as a waste of precious financial resources and therefore reduce firm value.

A third view suggests that there is no clear positive or negative relationship between the CSP and the CFP of a firm. Among others, the recent studies of Halbritter and Dorfleitner (2015) and Auer and Schuhmacher (2016) indicate that there is no statistical difference in the risk-adjusted returns of a portfolio consisting of either high ESG-rated or low ESG-rated firms. This third point of view does not necessarily conclude the absence of a connection between CSP and CFP but may, in contrast, on the one hand, indicate that the market prices CSP properly which leads to an absence of risk-adjusted returns, or, on the other hand, that the benefits resulting from the ESG-based activities will be offset by their respective drawbacks such as, for example, their costs or the occurrence of agency problems.

Whatever the relationship between CFP and CSP reveals itself to be in a specific context, the question of informational efficient markets still arises. As the stock selection of corresponding investment strategies is frequently based on the evaluation of certain ESG-based ratings, one may argue, as these scores are publicly available, that financially motivated investors could not generate a risk-adjusted excess return over conventional or non-ESG-based investments, due to of market efficiency. Fama (1965, 1970) describes, with the efficient market hypothesis (EMH), a framework in which, if the semi-strong form holds true, all information regarding the CSR of a company such as sustainability reports, ESG ratings, and even ESG-based scandals, should be correctly incorporated into the price of the respective stock shortly after being made public. Therefore, an outperformance of an ESG-based stock selection strategy would not be possible. However, Grossman (1976), and Grossman and Stiglitz (1980), for example, argue that a perfect information-efficient market could not exist, as there would be no incentive for investors to gather information or to actively manage a portfolio whatsoever, because they could not generate any excess returns.

In the case of SRI, Mynhardt et al. (2017) examine the efficiency of socially responsible indices by calculating a Hurst coefficient. The results indicate that most socially responsible indices are significantly less efficient than conventional ones. With a few exceptions, the Hurst coefficient of most of these indices differs from an efficient market (where the Hurst coefficient would be exactly 0.5), ranging either from 0.3 to 0.45 (signaling fat tails with an anti-persistent return series which is negatively correlated) or from 0.55 to 0.6 (indicating fat tails with a tendency to persistent return series with a slight positive correlation), which raises the question of whether ESG-based information is priced immediately and correctly and is considered in its entirety. This appears to be especially crucial in terms of ESG-based scandals as, whereas the occurrence of a scandal is publicly perceived and indeed undoubtedly immediately priced, the impact of the absence of these scandals has often been overlooked as companies with a low amount of scandals "fly under the radar". In this regard, the controversy score represents a good opportunity to decrease this ineffi-

ciency and can add significant value to ESG investing as this score is comparable to credit default ratings as these ratings also evaluate the absence of an infrequent event. Dorfleitner et al. (2018) also address the aspect of information inefficiency in the context of SRI as they argue that the future financial benefits of socially responsible activities are not immediately perceivable and therefore the economic nature of CSR remains fairly opaque. Within their results, they conclude that ESG-based activities lead to significant earnings surprises and unexpected additional cashflows in the long run. Edmans (2011) proves something similar with respect to the intangible asset of being one of the best companies to work for, due to the particularly good of their employees.

2.3 Data and methodology

2.3.1 Data

Due to their transparent scoring methodology we choose Thomson Reuters[1] as the world's largest ESG rating database for our data source (see, i.e., Cheng et al., 2014; Durand and Jacqueminet, 2015). Therefore, our dataset includes all Thomson Reuters scores (in the following referred to as TR scores), controversies and combined scores for the European, US, as well as the global market (including the US and European market) in the period under review from 2002 to 2018. These three scores represent the starting point for further calculations and are explained in more detail below.

First, the controversies scores, which pertain to Thomson Reuter's latest scoring methodology, add a new dimension to previous approaches by capturing negative media stories from global media sources. This score is a percentile ranking that takes ESG-based scandals into account concerning and infringing on any of the following controversy topics and that occur during a company's fiscal year. Its rating methodology consists of 23 ESG controversy topics such as "controversies privacy" or "business ethics controversies" (see Thomson Reuters, 2019). This score is also benchmarked on the respective industry groups.

Thus if a scandal occurs, it has a negative impact on the evaluation of the company involved. Ongoing legislation disputes, lawsuits, and fines may also affect the ensuing years and may still be visible in further controversy ratings. Furthermore, the valuation is as follows:

$$\text{score} = \frac{\#\text{ comp. with a worse value} + \frac{\#\text{ comp. with the same value included current one}}{2}}{\#\text{ comp. with a value}} \tag{2.1}$$

[1]The scores are currently published by Refinitiv.

In brief: the fewer scandals that affect a company, the higher its score is[2].

The TR score evaluates a company's environmental, social and corporate governance performance (ESG) with regard to ten main categories based on publicly available company-reported data. Each of these categories (for instance, resource use, innovation, and emissions in the environmental pillar, human rights, and workforce in the social pillar, and management in the corporate governance pillar) receives an individually calculated category score and a related category weighting within its associated pillar. These data result in three so-called pillar scores, one for each ESG pillar. To calculate the overall ESG score, these pillar scores are aggregated[3] and in the last step, the TR score is ranked by percentile and benchmarked against the industry. Therefore, the TR score implies an easy way to implement a best-in-class approach (see Thomson Reuters, 2019).

Next, the combined score comprises both the TR and the controversies score and thus offers a broadly diversified scoring with regard to performance-based ESG data and controversies collected from worldwide media sources (see Thomson Reuters, 2019). The controversies score has no impact on the TR score if it is greater than or equal to 50. In this case, the combined score equals the TR score. However, if the TR score is less than the controversies score, the combined score also equals the TR score. Only if the TR score is greater than the controversies score (< 50), the combined score equals the average of both scores[4].

In order to determine our data universe, we only consider companies for which all three ratings are present. Moreover, penny stocks are deleted. As a result, we obtain a monthly-based dataset with over 529,000 observations in total at an average of approximately 2,500 companies in a single month during our time period of 2002–2018 (192 months), more precisely between 900 and 4,700 at each point in time. For all observed companies, we have a comparable dataset of the three ratings (TR, combined and controversies). Table 2.1 shows the descriptive statistics of our data universe.

Table 2.1: Descriptive statistics.

Score	Mean	SD	Min	Max
TR	50.58	16.86	5.16	97.51
Controversies	49.49	20.27	0.08	90.91
Combined	45.46	15.51	5.16	95.22

This table presents the mean, standard deviation, minimum, and maximum values of the TR, controversies, and combined scores of the full dataset.

Concerning the TR rating, the mean value of the rating universe corresponds almost

[2]For more detailed information on the calculation, see Thomson Reuters (2019).

[3]The weightings of the three pillars are 34% for the environmental, 35.5% for the social and 30.5% for the governance pillar.

[4]For more detailed information on the calculation, see Thomson Reuters (2019).

exactly to 50 with a standard deviation of approximately 17. The controversies score is approximately the same as the TR score in terms of mean value and standard deviation. As can be expected with regard to the calculation, the combined score has a lower mean value than the TR and controversies score with a standard deviation of 15.

Regarding the correlation between the three scores it is noteworthy that the correlation between the controversies score and the TR score is negative (-0.3107). Thus, companies with a high TR score tend to have a low controversies score.

One explanation for this may be that companies that tend to have high ESG scores are affected more greatly by controversies, as reflected by the saying "the higher you fly, the harder you fall".

Furthermore, as would be expected from the composition, the correlation between TR score and combined score is positive (0.7774) as well as between controversies score and combined score (0.3077).

The analysis in this paper is carried out from the perspective of an US investor, so all data is converted into US dollars. The total returns and market capitalization of the considered companies are received from Thomson Reuters Eikon. Discarded (delisted) or insolvent companies are considered until the last available rating or financial information. Thus, our results are not influenced by a potential survivorship bias. For more detailed insights, some descriptives for the European and US market are displayed in Table 2.2. While for the European market we consider over 158,000 observations based on an average of approximately 820 companies (between 400 and 1,000), for the US market, our data consist of over 191,000 observations at an average of approximately 1,000 companies (between 400 and 2,300).

Table 2.2: Descriptive statistics for the European and US market.

	Europe			USA		
Score	Mean	SD	Observations	Mean	SD	Observations
TR	56.64	15.99	158,248	48.15	16.05	191,661
Controversies	48.36	21.24	158,248	46.53	21.91	191,661
Combined	50.30	15.50	158,248	42.08	14.03	191,661

This table presents the mean, standard deviation, and number of observations of the TR, controversies, and combined scores of the European and US datasets.

2.3.2 Methodology

As a first step, we construct several portfolios by generally sorting stocks according to each score. To calculate the monthly returns, we select the best-rated and worst-rated stocks, respectively, and combine them in a portfolio, one being for each of the three

scores. Following this procedure, we consider a best-only and worst-only strategy as well as a best-minus-worst strategy, which is long in the best-performing companies and short in the worst-performing ones. As a next step, we consider three different weighting approaches upon which to construct the portfolios. We include the common value-weighted and equally-weighted strategies and also a rank-weighted strategy that we present in detail below in Section 2.3.3.

We obtain nine stock portfolios[5] for value- and equally-weighted and rank-weighted strategies, which is the object of contemplation in Section 2.4.3, respectively in the European, US and global market – in total 27 per market. In order to determine the performance of our portfolios, we apply the Fama and French (2015) five-factor model, which is based on the regression:

$$\begin{aligned} R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t \\ + r_i RMW_t + c_i CMA_t + e_{it}. \end{aligned} \tag{2.2}$$

In this model, the return of portfolio i for period t is represented by R_{it} while R_{Ft} comprises the risk-free return. R_{Mt} denotes the return of the market portfolio, SMB_t represents the small-minus-big factor (returns of small stocks minus returns of big stocks) and HML_t is the performance difference between companies with a high and low book-to-market value. The factor RMW_t indicates the difference between the returns of stocks with a weak and a robust profitability. CMA_t describes the returns of conservative (i.e., low-investment firms) minus aggressive (i.e., high-investment firms) stocks. Moreover, b_i, s_i, h_i, r_i, and c_i are the estimated regression coefficients which are calculated by OLS regression, in which e_{it} denotes a (zero-mean) residual and a_i the intercept.

Since a Breusch and Pagan (1979) test applied to all portfolios indicates that the residuals of the regressions are subject of heteroskedasticity and a Godfrey (1978) and Breusch (1978) test as well as a Durbin and Watson (1971) test show autocorrelations for most of the models, we use the approach of Newey and West (1987) to calculate standard errors.

2.3.3 A different approach: rank-weighted portfolios

Besides equally-weighted and value-weighted portfolios, we also consider a new portfolio composition strategy following a similar approach to Frazzini and Pedersen (2014) which reflects the great importance of the ESG ratings for those investors, who may wish to award a different level in the scores through a corresponding weight. Consequently, we build portfolio weights based on the respective score placements. Our new approach is to award better scores and to consequently include them with higher weights in a best-portfolio strategy and vice versa in order to reward worse scores with higher weights in the

[5]This results from three different scores and three different portfolio sets.

worst portfolio. In addition, the best portfolios constructed this way have, by definition, a higher ESG rating than value-weighted or equally-weighted strategies, whereas the worst portfolios have lower ratings. First, we determine the best and worst stocks. Next, we divide the companies up by rank in ascending and descending order. In the best portfolios, the company with the highest score receives the (numerically) highest rank. In contrast, the company with the worst score receives the highest rank in the worst portfolios. To calculate the weights $w_{i,t}$ of a company $c \in C_t \subseteq C$, where C is the set of all companies within the respective data and C_t is the set of all companies within the portfolio at time t, we use

$$\begin{aligned} w_t \colon C_t \times T &\longrightarrow [0,1] \\ (c,t) &\longmapsto w_t(c,t) = \frac{(N_t - Rk_t(c)) + 1}{\sum_{\tilde{c} \in C_t} Rk_t(\tilde{c})} \end{aligned}$$

and for each $t \in T$ there holds

$$\sum_{\tilde{c} \in C_t} w_t(\tilde{c}, t) = 1,$$

where $Rk_t(c)$ note the rank of a company c at t, $N_t = |C_t|$ the cardinality of the portfolio selection at t, in the monthly period under review. If a company $\hat{c} \in C \backslash C_t$ does not appear in the portfolio selection at time t by definition, its weight is

$$w_t(\hat{c}, t) \coloneqq 0.$$

2.4 Results

2.4.1 Equally- and value-weighted portfolios

Table 2.3 presents some measures of all 27 equally-weighted 10% portfolio strategies. Concerning the Sharpe ratio, the Sortino ratio, and the Treynor ratio, it is noteworthy that all controversies best and TR worst portfolios show higher values than the respective market portfolio, which is a first indication that the performance of these portfolios is high. Furthermore, most best and worst portfolios have a higher risk than their respective market in terms of maximum drawdown (MDD), while the controversies best-minus-worst portfolios have a much lower risk in all three markets. Additionally, the MDD is lower than that of the corresponding market for the following portfolios: combined best-minus-worst (US, global), controversies best (Europe, global), TR worst (global) and combined worst (European).

Table 2.3: Measures for equally-weighted 10% portfolios.

		MDD	Skewness	Kurtosis	Sharpe ratio	Sortino ratio	Treynor ratio
Europe							
TR	Best	0.6245	−0.3056	1.5780	0.3476	0.1939	0.0687
	Worst	0.6387	−0.3815	1.8294	0.6442	0.3244	0.1287
	Best-worst	0.6213	−0.2553	0.8853	−0.9329	−0.3059	−1.9748
Controversies	Best	0.5696	−0.4338	2.0089	0.6817	0.3352	0.1363
	Worst	0.6414	−0.2846	1.9547	0.2721	0.1636	0.0542
	Best-worst	0.1652	−0.6429	2.6383	0.4591	0.2070	−0.1891
Combined	Best	0.6213	−0.5126	2.0035	0.3854	0.2065	0.0760
	Worst	0.5696	−0.4055	2.1586	0.4544	0.2401	0.0916
	Best-worst	0.6414	−0.3854	0.8504	−0.4932	−0.1628	0.9960
Market		0.5903	−0.6565	1.7155	0.3521	0.1818	0.0650
USA							
TR	Best	0.5112	−0.3836	3.1051	0.4932	0.2452	0.0787
	Worst	0.5119	−0.2851	2.5200	0.6032	0.3043	0.0985
	Best-worst	0.5458	−0.5538	2.6846	−0.7360	−0.2448	0.3503
Controversies	Best	0.5320	−0.1696	1.8906	0.6769	0.3428	0.1121
	Worst	0.5571	−0.1891	2.8373	0.4623	0.2403	0.0745
	Best-worst	0.1529	0.5305	1.6073	0.3448	0.1798	−2.3288
Combined	Best	0.5137	−0.2660	3.4409	0.5532	0.2790	0.0899
	Worst	0.5710	−0.2159	2.6147	0.5870	0.2987	0.0957
	Best-worst	0.3712	−0.1706	2.2610	−0.4897	−0.1711	0.3478
Market		0.5039	−0.6927	1.6337	0.4783	0.2238	0.0687
Global							
TR	Best	0.5591	−0.4751	2.3367	0.4776	0.2395	0.0760
	Worst	0.5259	−0.3081	2.5068	0.7753	0.3793	0.1266
	Best-worst	0.6416	−0.7139	2.5614	−1.0850	−0.3442	−8.5105
Controversies	Best	0.5136	−0.4776	2.3364	0.7892	0.3774	0.1273
	Worst	0.6084	−0.4332	2.5499	0.3906	0.2038	0.0631
	Best-worst	0.1201	0.1383	1.3355	0.4274	0.2139	−0.1457
Combined	Best	0.5676	−0.5478	2.9513	0.5282	0.2598	0.0840
	Worst	0.5637	−0.2991	2.6915	0.6707	0.3313	0.1094
	Best-worst	0.4704	−0.7383	3.1132	−0.7641	−0.2549	2.3499
Market		0.5363	−0.8494	2.4412	0.4457	0.2094	0.0670

This table shows the maximum drawdown (MDD), skewness, kurtosis (excess), Sharpe ratio, Sortino ratio, and Treynor ratio for portfolios from 2002 to 2018. The variables are calculated individually for each equally-weighted portfolio based on a 10% cutoff of each score, market, and portfolio set as well as for the respective total market.

Table 2.4: Equally-weighted 10% portfolios: regressions based on the three observed markets.

		Alpha	MKT	SMB	HML	RMW	CMA	Adj. R^2
Europe								
TR	Best	0.0016	1.0530***	−0.2484***	0.3284**	−0.0451	−0.1075	0.8843
	Worst	0.0048***	1.0084***	0.4390***	0.3637**	−0.1765	−0.2130	0.8831
	Best-worst	−0.0043***	0.0464	−0.6822***	−0.0463	0.1297	0.1104	0.3869
Controversies	Best	0.0049***	0.9926***	0.2184***	−0.0461	−0.0545	0.0385	0.8546
	Worst	0.0020	1.0651	−0.0946	0.2418	−0.3835*	−0.3656*	0.8699
	Best-worst	0.0020	−0.0706**	0.3181***	−0.2990**	0.3272*	0.4090***	0.3312
Combined	Best	0.0004	1.0816***	−0.0010	0.3257**	0.1577	−0.0421	0.8845
	Worst	0.0035*	1.0448***	0.3742***	0.2493	−0.3092*	−0.4445*	0.8567
	Best-worst	−0.0040**	0.0387	−0.3700***	0.0654	0.4652***	0.4073**	0.1987
USA								
TR	Best	0.0017	1.0964***	0.0786	0.2128***	−0.0770	−0.2334	0.8192
	Worst	0.0044***	1.1448***	0.5243***	0.3307***	−0.0966	−0.5281***	0.8341
	Best-worst	−0.0037***	−0.0446	−0.4440***	−0.1228**	0.0231	0.2694	0.3687
Controversies	Best	0.0049**	1.0739***	0.4290***	0.1796*	−0.1431	−0.3313*	0.7881
	Worst	0.0019	1.1495***	0.1577*	0.3112***	−0.0437	−0.3660	0.8151
	Best-worst	0.0020	−0.0718	0.2730***	−0.1364*	−0.0963	0.0390	0.0828
Combined	Best	0.0019	1.1568***	0.3669***	0.3015***	0.1347	−0.3785*	0.8233
	Worst	0.0045***	1.1546***	0.4696***	0.3720***	−0.1067	−0.6109***	0.8341
	Best-worst	−0.0035***	0.0060	−0.1009	−0.0754	0.2445***	0.2370**	0.1165
Global								
TR	Best	0.0025*	1.1249***	−0.0934	0.1997**	−0.3808***	−0.0858	0.9011
	Worst	0.0071***	1.0647***	0.3075***	0.1991	−0.4761***	−0.4157	0.8664
	Best-worst	−0.0056***	0.0633**	−0.3911***	−0.0117	0.1004	0.3381	0.1669
Controversies	Best	0.0056***	0.9958***	0.2406***	0.1073	−0.3080***	−0.1590	0.8772
	Worst	0.0022	1.1654***	−0.0865	0.2515	−0.4790***	−0.2810	0.8737
	Best-worst	0.0023*	−0.1666***	0.3369***	−0.1565	0.1761	0.1302	0.2535
Combined	Best	0.0020	1.1496***	0.0970	0.2779***	−0.1339	−0.1500	0.9010
	Worst	0.0061***	1.0859***	0.2420***	0.2294*	−0.4990***	−0.4516*	0.8645
	Best-worst	−0.0051***	0.0668**	−0.1353*	0.0361	0.3701***	0.3099***	0.0887

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each equally-weighted portfolio based on a 10% cutoff of each score, market, and portfolio set. The best (worst) portfolios consist of the 10% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas, all estimated coefficients of the five Fama and French (2015) factors and adj. R^2 are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

Table 2.5: Value-weighted 10% portfolios: regressions based on the three observed markets.

		Alpha	MKT	SMB	HML	RMW	CMA	Adj. R^2
Europe								
TR	Best	−0.0028**	1.0161***	−0.3915***	0.3611***	0.2073	−0.1933	0.8690
	Worst	−0.0037	0.8806***	0.4401*	0.2633	−0.1965	−0.0173	0.4343
	Best-worst	−0.0001	0.1373	−0.8265***	0.0868	0.4020	−0.1711	0.0684
Controversies	Best	0.0019	0.9588***	−0.0932	−0.0545	0.4038***	0.2531*	0.7984
	Worst	−0.0015	1.0022***	−0.3708***	0.2838**	0.2134	−0.0585	0.8736
	Best-worst	0.0024	−0.0415	0.2827***	−0.3493***	0.1885	0.3165**	0.1851
Combined	Best	−0.0040**	1.0891***	−0.1922*	0.2496*	0.2226	0.0047	0.8218
	Worst	−0.0050	0.9880***	0.2460	0.0022	−0.3334	−0.1350	0.5185
	Best-worst	0.0000	0.1030	−0.4331*	0.2364	0.5543	0.1445	0.0108
USA								
TR	Best	−0.0002	0.9763***	−0.2039***	0.2189***	−0.0133	−0.2165**	0.8199
	Worst	0.0016	1.0369***	0.1796**	0.1176	−0.0991	−0.5176***	0.7567
	Best-worst	−0.0028	−0.0569	−0.3817***	0.0965	0.0889	0.3057**	0.1604
Controversies	Best	0.0017	0.9489***	0.1649**	0.1628*	0.0511	−0.3559***	0.7433
	Worst	−0.0006	1.0348***	−0.2573***	0.2047***	−0.0140	−0.2468*	0.8289
	Best-worst	0.0012	−0.0821	0.4239***	−0.0467	0.0683	−0.1044	0.1300
Combined	Best	0.0003	1.0341***	0.1473**	0.1858**	0.1454	−0.3361***	0.8147
	Worst	0.0016	1.1014***	0.0365	0.0755	−0.1457	−0.5575***	0.7864
	Best-worst	−0.0023	−0.0635	0.1126	0.1055	0.2942***	0.2260*	0.0900
Global								
TR	Best	−0.0009	1.0247***	−0.3855***	0.2376***	−0.1266	−0.1252	0.8919
	Worst	0.0013	0.9584***	0.0214***	−0.1020	−0.0148	−0.3126*	0.7645
	Best-worst	−0.0033*	0.0694	−0.3971***	0.3273***	−0.1067	0.1957	0.1411
Controversies	Best	0.0033***	0.8916***	0.0007	−0.0719	0.1870	−0.1538	0.7969
	Worst	0.0002	1.0422***	−0.4235***	0.1572*	−0.1315	−0.1250	0.8915
	Best-worst	−0.0021	−0.1474***	0.4340***	−0.2414**	0.3236**	−0.0205	0.2313
Combined	Best	−0.0004	1.0376***	−0.0670	0.1147	−0.0622	−0.2059*	0.8761
	Worst	0.0015	1.0261***	−0.0456	−0.2160*	−0.1146	−0.3610**	0.8254
	Best-worst	−0.0030*	0.0145	−0.0116	0.3183***	0.0575	0.1633	0.0734

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each value-weighted portfolio based on a 10% cutoff of each score, market, and portfolio set. The best (worst) portfolios consist of the 10% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas, all estimated coefficients of the five Fama and French (2015) factors and adj. R^2 are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

To examine a potential over-performance of the strategies in more detail, we consider the alphas of the respective portfolios. The results of the Fama and French (2015) five-factor regressions are presented in Table 2.4 for equally-weighted portfolios and in Table 2.5 for value-weighted portfolios. Some results immediately catch the eye: Regarding the equally-weighted strategy, the worst portfolios based on the TR and combined scores, as well as the best portfolios of the controversies score, indicate positive and significant outperformance. For the controversies score best portfolios, consistently positive and significant alphas can be observed for all portfolios. These portfolios show strongly significant returns of up to almost 7% p.a.[6]. In contrast to this, the controversies score worst and best-minus-worst portfolios do not exhibit any striking features.

Surprisingly, when considering combined score portfolios, a best portfolio strategy does not lead to a significant performance. However, the performance of the worst portfolio shows a consistently strong and significant outperformance of up to about 7.6% p.a., which can be observed in all three markets. As a result of this, the calculations indicate a significant underperformance of the best-minus-worst portfolios. Therefore, this effect cannot be caused by the controversies score, but instead appears to be determined by the second component of the combined score, namely the TR score.

When taking a closer look at the ESG portfolios, we notice the following. While the performance of the best portfolios—apart from a slight significance in the global market—does not show any over-performance, a strongly significant outperformance of up to almost 9% (8.86%) p.a. can be observed for the worst TR score portfolios in all three markets. These results resemble those of the combined score portfolios.

On the contrary, we compare this with the results of the value-weighted portfolios in Table 2.5. Apart from very few exceptions neither best nor worst portfolios based on the three ratings obtain any ongoing positive and significant alphas within the European, US or global market. So, it becomes relatively clear that there are no ongoing tendencies recognizable in terms of any benefits of best or worst strategies. Apart from some isolated outliers, the results lead us to the assumption that the value-weighted strategy does not result in any excess return for investors, which is consistent with the findings of Halbritter and Dorfleitner (2015). It should also be pointed out that the adjusted R^2 values of all long and short portfolios are consistently high, which indicates a strong explanatory power of our underlying factor model.

There is a clearly recognizable difference between Tables 2.4 and 2.5: Since the results of the value-weighted and the equally-weighted portfolios are very distinct, this points to the fact that the significant outperformance of the equally-weighted portfolios is strongly driven by the small companies. In particular, the TR portfolios support the above finding

[6]The annualized performance of the global controversies score best portfolio is: $1.0056^{12} - 1 = 0.0693$.

as the equally-weighted portfolios based on low TR scores achieve strong outperformance. These results provide some evidence of the trade-off hypothesis (see Aupperle et al., 1985), as investors appear to reward smaller companies for not investing their money in ESG improvements. They may consider this spending as a wasteful investment and prefer companies that invest in growth and innovation. As no or even negative significant results were shown for value-weighted best portfolios, we can conclude that, for large companies, the benefits of expenditures improving CSP are already reflected in the stock price of these companies.

Looking at the data, it becomes apparent that an equally-weighted portfolio strategy based on a high controversies score leads to a high outperformance. Therefore, this demonstrates that small companies in particular generate a sustained stock performance if they have a "clean coat" with regard to controversies. Thus, one might say that they "fly under the radar".

Last but not least, the above observations also find their reflection in the combined score portfolios. On the one hand, the effect of the TR worst portfolios also occurs in the combined score worst portfolios, which are by definition strongly influenced by the TR score. On the other hand, it is not surprising that a slight decrease in the returns appears in these portfolios compared with corresponding TR worst portfolios, which can be explained due to the influence of the controversies score.

To discuss these results against the background of current literature, it is necessary to divide this step into two parts. As already published by previous studies such as Halbritter and Dorfleitner (2015), we confirm the recent observation, being that a market-weighted ESG strategy does not result in ongoing significant overperformance, so for this strategy, there is no clear out- or underperformance of best or worst portfolios.

The hypothesis of a positive relationship between the CSP and the CFP of a company (see, e.g., Kempf and Osthoff, 2007) could only partly be confirmed. Evidently, there is no performance loss when investing in ESG portfolios, but the data suggest that there is also no ongoing positive outperformance for companies with high ESG ratings, so for these portfolios, we strongly support the results of Revelli and Viviani (2015), being that neither weaknesses nor strengths can be detected for value-weighted positive CSP strategies.

However, this is reverted when considering equally-weighted portfolios. Remarkably, no significant negative performance is detected when investing in best ESG portfolios with an equally-weighted strategy. Thus, there are no ESG-based performance losses for investors. Moreover, Statman and Glushkov (2009) find that investors can achieve positive abnormal returns with socially responsible top-minus-bottom strategies using equally-weighted portfolios. Thus, in relation to the results of our best-worst portfolios, there is no reason for investors to pursue this strategy nowadays because, in particular, the worst portfolios

based on the TR score reveal a significant overperformance. However, this also stands in contradiction to Auer (2016), who claims that investors should eliminate firms with the worst ESG ratings, whereas we find evidence of the fact that these represent some potential for (ESG neutral) investors. Moreover, this finding contradicts even Kempf and Osthoff (2007), who use a long-short strategy and obtain an overperformance. Contrary to this and related to our results, doing good while doing well did not manifest itself at all during our work.

Market efficientists would expect an immediate reaction on the stock market in the face of a controversy. Therefore, no long-term overperformance can be expected with regard to market-efficiency aspects, so it is surprising that there are several corresponding findings for the controversies score portfolios. Although the occurrences of controversies may be immediately priced by the market, which is indicated by the non-existing underperformance of the worst controversies score portfolio, the absence of controversies appears to be incorrectly evaluated for small companies. The significant outperformance of the best-rated companies therefore indicates a less efficient market regarding ESG-based information as discussed by Edmans (2011), Mynhardt et al. (2017), and Dorfleitner et al. (2018). Smaller companies without an unwanted boost in public perception due to a controversy remain "silent saints" so-to-speak, and "fly under the radar". The controversies score enables a valuation of controversies that do *not* take place and may therefore be a good tool to enhance ESG investment as it reveals companies with a low amount of scandals with a specific potential for an increase in market value and stock price.

An additional consideration of the Fama and French factor coefficients yields some interesting insights regarding the differences between value and equally weighting. First, it can be seen that the market betas are generally around 1, but tend to be lower for value-weighted portfolios. This is not surprising, as smaller companies may have higher market betas and these companies are represented with higher weigths in the equally-weighted portfolios. Second, we notice that the controversies best, TR worst, and combined worst equally-weighted portfolios have significant positive SMB_t factor coefficients and reveal a higher absolute value compared to the respective value-weighted portfolios, which is again explainable by the higher weights for smaller companies. Third, the remaining factors show no systematically deviating patterns.

2.4.2 Portfolios based on market capitalization

To further investigate whether the observed strong overperformance of equally-weighted portfolios with low TR ratings and high controversies scores is driven by company size, we divide our dataset at the median of the market capitalization and create new portfolios based on companies with high and low market capitalizations. Table 2.6 displays these

portfolios based on a 10% cutoff for the European, US and global markets. From this table, it is apparent that the main results remain consistent, namely a significant outperformance of portfolios based on small companies with low TR score ratings as well as portfolios based on small companies with fewer controversies and therefore high controversies score. It also can be seen from Table 2.6 that even the value-weighted calculations based on firms with low market capitalization mostly show significant and positive alphas for controversies best, TR worst portfolios and ensure our results.

2.4.3 Rank-weighted portfolios

Table 2.7 displays best and worst rank-weighted portfolios based on a 10% cutoff for the European, US and global market. When considering these portfolios, nearly all returns of the best and worst portfolios are higher than with the corresponding equally-weighted strategies. Based on these calculations, the returns improve by up to 42.86%[7] for the best, by up to 32.24%[8] for the worst and by up to 84.28%[9] for the best-minus-worst portfolios, compared with the corresponding equally-weighted portfolios. Note that rank-weighted portfolios also reveal a lower significance level in terms of p-values, which indicates a real potential for investors.

On the one hand, there are a number of promising investment strategies for investors who strongly attach importance to ESG scores. As we previously mentioned, the controversies score represents a huge potential for investors in particular, and together with a rank-weighted portfolio strategy the corresponding alphas even increase, so this score describes a way in which to detect companies with a specific management culture that apparently leads to higher future cash flows and therefore to higher and more significant alphas. Surprisingly, companies with a high controversies score do not necessarily have a high ESG score. This noteworthy observation remains open for future research.

On the other hand, investors pursuing exactly the opposite strategy also benefit from rank weighting portfolios. This is particularly evident in the outperformance of the TR worst portfolios. Obviously, stronger weightings for firms with very low TR scores lead to significant overperformance, which can be traced back to a trade-off interpretation (see Aupperle et al., 1985). In summary, one can conclude that the rank weighting portfolios represent a useful tool for investors who wish to profit from ESG ratings either by investing in high ranked companies or by investing in low-ranked firms. Finally, to put it in a nutshell: buy the "saints" or invest in the "small sinners".

[7]This displays the improvement in annual returns from 0.0693 to 0.0990 of the global controversies best portfolio.

[8]This displays the improvement in annual returns from 0.0428 to 0.0566 of the Europe combined worst portfolio.

[9]This displays the improvement in annual returns from 0.0280 to 0.0516 of the global controversies best-worst portfolio.

Table 2.6: Alphas of equally- and value-weighted 10% portfolios: regression based on high and low market capitalization.

		Europe				USA				Global			
		High MC		Low MC		High MC		Low MC		High MC		Low MC	
		EW	VW	EW	VW	EW	VW	EW	VW	EW	VW	EW	VW
TR	Best	0.0014	−0.0017	0.0032	0.0021	0.0023*	0.0007	0.0050*	0.0035	0.0013	−0.0016	0.0058**	0.0046**
	Worst	0.0003	−0.0043	0.0079***	0.0047***	0.0005	−0.0013	0.0064***	0.0034*	0.0016	−0.0011	0.0107***	0.0057***
	Best-worst	0.0001	0.0016	−0.0057***	−0.0036*	0.0008	0.0009	−0.0024	−0.0009	−0.0013	−0.0016	−0.0059***	−0.0021*
Controversies	Best	0.0020	0.0011	0.0078***	0.0061***	0.0007	0.0008	0.0077***	0.0040*	0.0022**	0.0029*	0.0087***	0.0061***
	Worst	0.0024	−0.0011	0.0034	0.0007	0.0013	−0.0005	0.0031	0.0012	0.0023*	−0.0002	0.0033	0.0019
	Best-worst	−0.0014	0.0012	0.0034	0.0044	−0.0016	0.0002	0.0035	0.0018	−0.0011	0.0020	0.0044**	0.0032*
Combined	Best	0.0001	−0.0047**	0.0025	0.0021	0.0014	0.0005	0.0047**	0.0030	−0.0002	−0.0020	0.0053**	0.0046**
	Worst	−0.0007	−0.0025	0.0055**	0.0036*	0.0010	0.0015	0.0050*	0.0023	0.0016	−0.0008	0.0085***	0.0042***
	Best-worst	−0.0002	−0.0032	−0.0039**	−0.0025	−0.0006	−0.0021	−0.0013	−0.0003	−0.0028**	−0.0022	−0.0042**	−0.0007

This table shows the alphas of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each equally- and value-weighted portfolio based on a 10% cutoff of each score, market, and portfolio. The calculations are performed on the basis of our dataset divided by the median of the market capitalization. The best (worst) portfolios consist of the 10% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

Table 2.7: Rank-weighted 10% portfolios: regressions based on the three observed markets.

		Europe		USA		Global	
		Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2
TR	Best	0.0020	0.8718	0.0017	0.7899	0.0022	0.8918
	Worst	0.0057***	0.8716	0.0047**	0.8189	0.0077***	0.8502
	Best-worst	−0.0047***	0.3329	−0.0040***	0.3344	−0.0065***	0.1627
Controversies	Best	0.0064***	0.8575	0.0062***	0.7957	0.0079***	0.8777
	Worst	0.0031*	0.8541	0.0014	0.8216	0.0027	0.8667
	Best-worst	0.0023	0.3216	0.0038**	0.2049	0.0042***	0.2429
Combined	Best	0.0010	0.8761	0.0017	0.8177	0.0018	0.9021
	Worst	0.0046**	0.8436	0.0040*	0.8163	0.0063***	0.8515
	Best-worst	−0.0045**	0.2231	−0.0033***	0.1522	−0.0056***	0.0992

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each rank-weighted portfolio based on a 10% cutoff of each score, market, and portfolio set. The best (worst) portfolios consist of the 10% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas and adj. R^2 are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

2.5 Robustness checks

To check our results for robustness, we run some further regressions. First of all, we construct the equally-weighted portfolios based on the 20% (instead of 10%) best and worst companies. Again we use the Fama and French (2015) five-factor regression model. The results are presented in Table 2.8 and indicate that all previous results remain materially the same for the 20% equally-weighted selection, i.e., an outperformance of the controversies score best and the TR and combined score worst portfolios.

Moreover, with regard to the rank-weighted strategy, the 20% portfolios are also examined. Following the same procedure, this leads to the results displayed in Table 2.9. Also, in this case, all results of previous calculations remain approximately unchanged. Compared with the 20% equally-weighted portfolios, most of the alphas are higher. For instance, we can observe an almost 20% increase in the alpha of the controversies best portfolio in the global market from 0.0046 to 0.0055, both being significant at a 1% level.

As a next step, we divide our portfolios into bull and bear market periods to monitor how the portfolio strategies perform in different market phases. The results are shown in Table 2.10. The data suggest that the majority of the strategies work in bull markets. Moreover, one argument against this cannot be ignored: In our investigation period, there were mostly bullish phases and only a few bearish time periods, those of which are comparatively short. Since we are nevertheless also able to detect a number of positive significant results in bearish market phases, for example the best controversies portfolio in the US market or most portfolios in the global market, this points to the fact that the strategies are robust against various market movements.

Furthermore, we split our portfolios up into two subperiods (Table 2.11). The first subperiod dates from April 2002 to March 2010 and the second from April 2010 until April 2018. The findings show for the US and global portfolios in particular that the abnormal returns are maintained even under this sample split. Eventually, we also check the results for a winsorization of the returns at the 1% level and re-run all regressions. The results remained unchanged.

In addition, we also construct equally- and value-weighted portfolios based on 20% (instead of 10%) best and worst companies with high and low market capitalization. The results of these regressions are displayed in Table 2.12. All previous major results remain materially unchanged for the 20% portfolios.

In order to include transaction costs, it is necessary to account for the turnover rate of the considered portfolios. For the 10% cutoff and US portfolios, we observe an average monthly turnover of 6.74% for the best TR and 8.55% for the worst TR, respectively 11.82% and 9.15% for the controversies score, as well as 8.84% and 9.69% for the combined score

portfolios. This remains on an equal level for the other markets under review, so that the average monthly turnover rate stands at approximately 10%. Even for all other cutoffs, the turnover rate is materially the same. Thus, in line with Frazzini et al. (2018), the results of these portfolio strategies lead to expected annual trading costs between 90 and 150 bps, which implies that the significant alphas remain positive even after transaction costs.

Table 2.8: Equally-weighted 20% portfolios: regressions based on the three observed markets.

		Europe		USA		Global	
		Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2
TR	Best	0.0014	0.8956	0.0020	0.8297	0.0023*	0.9044
	Worst	0.0043**	0.8871	0.0041**	0.8549	0.0059***	0.8804
	Best-worst	−0.0039***	0.4758	−0.0031***	0.4006	−0.0046***	0.2167
Controversies	Best	0.0051***	0.8952	0.0038***	0.8102	0.0046***	0.8780
	Worst	0.0018	0.8854	0.0017	0.8279	0.0021	0.8854
	Best-worst	0.0024**	0.2740	0.0011	0.0577	0.0015	0.1443
Combined	Best	0.0020	0.8860	0.0023	0.8226	0.0026*	0.9103
	Worst	0.0030*	0.8864	0.0030**	0.8549	0.0053***	0.8850
	Best-worst	−0.0019*	0.2459	−0.0018**	0.0850	−0.0038***	0.0920

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each equally-weighted portfolio based on a 20% cutoff of each score, market, and portfolio set. The best (worst) portfolios consist of the 20% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas and adj. R^2 are reported upon. In oder to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

Table 2.9: Rank-weighted 20% portfolios: regressions based on the three observed markets.

		Europe		USA		Global	
		Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2
TR	Best	0.0018	0.8884	0.0016	0.8214	0.0024*	0.9012
	Worst	0.0049***	0.8853	0.0041**	0.8455	0.0069***	0.8728
	Best-worst	−0.0042***	0.4182	−0.0035***	0.4105	−0.0056***	0.1868
Controversies	Best	0.0053***	0.8838	0.0046**	0.8128	0.0057***	0.8789
	Worst	0.0022	0.8772	0.0016	0.8311	0.0022	0.8804
	Best-worst	0.0021*	0.3409	0.0020	0.1333	0.0025**	0.2523
Combined	Best	0.0015	0.8831	0.0022	0.8184	0.0023*	0.9056
	Worst	0.0041**	0.8738	0.0036**	0.8416	0.0058***	0.8729
	Best-worst	−0.0036***	0.2667	−0.0025***	0.1275	−0.0045***	0.1019

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each rank-weighted portfolio based on a 20% cutoff of each score, market, and portfolio set. The best (worst) portfolios consist of the 20% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas and adj. R^2 are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

Table 2.10: Bull and bear market portfolios.

		Europe				USA				Global			
		Alpha bull	Adj. R^2	Alpha bear	Adj. R^2	Alpha bull	Adj. R^2	Alpha bear	Adj. R^2	Alpha bull	Adj. R^2	Alpha bear	Adj. R^2
TR	Best	0.0003	0.8743	0.0005	0.8840	0.0000	0.8404	0.0186*	0.7629	0.0008	0.8943	0.0127**	0.9186
	Worst	0.0042**	0.8550	−0.0002	0.9132	0.0030**	0.8140	0.0148	0.8067	0.0051***	0.8276	0.0104**	0.9259
	Best-worst	−0.0048***	0.3944	−0.0006	0.4646	−0.0039***	0.4258	0.0022	0.2103	−0.0053***	0.2048	0.0009	0.3559
Controversies	Best	0.0048**	0.8129	0.0058	0.8750	0.0033*	0.7553	0.0151*	0.7963	0.0049***	0.8410	0.0107*	0.8914
	Worst	0.0003	0.8578	−0.0007	0.8780	0.0000	0.8214	0.0160	0.7485	0.0001	0.8666	0.0130**	0.8729
	Best-worst	0.0036**	0.3118	0.0052	0.5556	0.0023	0.1195	−0.0023	−0.0921	0.0037**	0.2072	−0.0037	0.6243
Combined	Best	0.0003	0.8755	−0.0004	0.8502	−0.0002	0.8349	0.0142	0.7520	0.0009	0.8842	0.0084**	0.9166
	Worst	0.0033	0.8174	−0.0051*	0.8999	0.0033**	0.8102	0.0143	0.8099	0.0044***	0.8242	0.0101**	0.9055
	Best-worst	−0.0039**	0.1933	0.0034	0.0940	−0.0044***	0.1514	−0.0015	0.0100	−0.0044***	0.0748	−0.0031	0.3928

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 divided into bull and bear market periods. The regressions are calculated individually for each equally-weighted portfolio based on each score, market, and portfolio set. The best (worst) portfolios consist of the best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas and adj. R^2 are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

Table 2.11: Subperiod portfolios.

			Europe				USA				Global			
			10%		20%		10%		20%		10%		20%	
			Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2	Alpha	Adj. R^2
TR	Best	4/02 - 3/10	0.0029	0.8772	0.0021	0.8896	0.0053*	0.8049	0.0056**	0.8156	0.0057**	0.8991	0.0051**	0.9044
		4/10 - 3/18	0.0004	0.8936	0.0009	0.9052	−0.0009	0.8655	−0.0010	0.8689	−0.0008	0.9145	−0.0005	0.9108
	Worst	4/02 - 3/10	0.0068***	0.8848	0.0066**	0.8886	0.0062**	0.8308	0.0065**	0.8418	0.0098***	0.8981	0.0089***	0.8909
		4/10 - 3/18	0.0029	0.8794	0.0019	0.8916	0.0040**	0.8453	0.0030*	0.8835	0.0043**	0.8232	0.0029*	0.8709
	Best-worst	4/02 - 3/10	−0.0057***	0.3237	−0.0064***	0.4403	−0.0028*	0.2938	−0.0027**	0.2843	−0.0060***	0.2234	−0.0057***	0.2416
		4/10 - 3/18	−0.0026	0.4556	−0.0011	0.5545	−0.0051***	0.5220	−0.0041***	0.6138	−0.0052***	0.2282	−0.0035***	0.2776
Controversies	Best	4/02 - 3/10	0.0076***	0.8817	0.0083***	0.9039	0.0096***	0.8036	0.0078***	0.8066	0.0084***	0.9047	0.0070***	0.9068
		4/10 - 3/18	0.0028	0.8273	0.0019	0.8957	0.0016	0.7811	0.0010	0.8302	0.0035**	0.8420	0.0023	0.8439
	Worst	4/02 - 3/10	0.0040	0.8762	0.0030	0.8895	0.0067**	0.8078	0.0056**	0.8233	0.0057**	0.8766	0.0053**	0.8903
		4/10 - 3/18	−0.0003	0.8710	0.0002	0.8889	−0.0019	0.8566	−0.0013	0.8446	−0.0014	0.8815	−0.0013	0.8895
	Best-worst	4/02 - 3/10	0.0017	0.3875	0.0034*	0.2931	0.0010	0.1029	−0.0014	0.0883	0.0008	0.3832	−0.0002	0.3262
		4/10 - 3/18	0.0029	0.2537	0.0015	0.2737	0.0033	0.1174	−0.0029***	0.2917	0.0047**	0.1271	0.0035*	0.0759
Combined	Best	4/02 - 3/10	0.0004	0.8746	0.0021	0.8819	0.0059**	0.8089	0.0063**	0.8089	0.0046**	0.9044	0.0058**	0.9113
		4/10 - 3/18	0.0065	0.8946	0.0015	0.8909	−0.0018	0.8559	−0.0010	0.8581	−0.0006	0.8984	−0.0006	0.9142
	Worst	4/02 - 3/10	0.0039	0.8806	0.0048*	0.8843	0.0067**	0.8289	0.0058**	0.8441	0.0088***	0.8905	0.0084***	0.8987
		4/10 - 3/18	0.0021	0.8226	0.0011	0.8954	0.0093**	0.8560	0.0018	0.8834	0.0033*	0.8335	0.0021	0.8737
	Best-worst	4/02 - 3/10	−0.0054***	0.1895	−0.0046**	0.1812	−0.0027*	0.1291	−0.0014	0.0883	−0.0060***	0.2206	−0.0045***	0.2128
		4/10 - 3/18	−0.0016	0.2368	0.0002	0.3785	−0.0059***	0.3071	−0.0029***	0.2917	−0.0041**	0.1694	−0.0029**	0.1757

This table shows the results of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis divided into two subperiods. The first subperiod dates from April 2002 to March 2010 and the second from April 2010 until April 2018. The regressions are calculated individually for each equally-weighted portfolio based on a 10% and 20% cutoff of each score, market, and portfolio set. The best (worst) portfolios consist of the 10% and 20% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas and adj. R^2 are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

Table 2.12: Alphas of equally- and value-weighted 20% portfolios: regression based on high and low market capitalization.

		Europe				USA				Global			
		High MC		Low MC		High MC		Low MC		High MC		Low MC	
		EW	VW	EW	VW	EW	VW	EW	VW	EW	VW	EW	VW
TR	Best	0.0010	−0.0026**	0.0029	0.0019	0.0013	−0.0003	0.0042	0.0028	0.0014	−0.0010	0.0048**	0.0035**
	Worst	0.0004	−0.0025	0.0057**	0.0027*	−0.0006	−0.0012	0.0062***	0.0037**	0.0008	−0.0002	0.0092***	0.0052***
	Best-worst	−0.0003	−0.0011	−0.0038***	−0.0017	0.0009	−0.0001	−0.0030**	−0.0019*	−0.0004	−0.0018	−0.0054***	−0.0027***
Controversies	Best	0.0032**	0.0014	0.0065***	0.0053***	−0.0004	−0.0003	0.0077***	0.0047**	0.0008	0.0014	0.0077***	0.0050***
	Worst	0.0005	−0.0008	0.0033	0.0014	0.0005	−0.0005	0.0055*	0.0028	0.0014	0.0002	0.0053**	0.0035*
	Best-worst	0.0017	0.0013	0.0022	0.0029	−0.0019	−0.0008	0.0012	0.0009	−0.0017	0.0003	0.0013	0.0005
Combined	Best	0.0006	−0.0029	0.0028	0.0023	0.0007	0.0002	0.0057**	0.0035*	0.0006	0.0011	0.0050**	0.0037**
	Worst	−0.0001	−0.0018	0.0050**	0.0027*	0.0001	−0.0007	0.0048**	0.0020	0.0014	0.0004	0.0078***	0.0043***
	Best-worst	−0.0003	−0.0021	−0.0032**	−0.0014	−0.0004	−0.0001	−0.0001	0.0005	−0.0018**	−0.0004	−0.0038***	−0.0017**

This table shows the alphas of the Fama and French (2015) five-factor regression for portfolios from 2002 to 2018 on a monthly basis. The regressions are calculated individually for each equally- and value-weighted portfolio based on a 20% cutoff of each score, market, and portfolio. The calculations are performed on the basis of our dataset divided by the median of the market capitalization. The best (worst) portfolios consist of the 20% best (worst) rated companies regarding a particular score. The best-worst portfolios are long in the best-performing companies and short in the worst-performing ones. Monthly alphas are reported upon. In order to estimate standard errors, we use the Newey and West (1987) procedure. ***,**, and * indicate a significance level of 1%, 5%, and 10%.

2.6 Conclusion

In this paper, we examine a dataset that includes over 4,700 companies and the associated TR, controversies, and combined scores in the Thomson Reuters Eikon universe in the investigation period from 2002 to 2018. All calculations are performed for the European, US and global markets. This paper is the first one investigating positive screened portfolios dependent on the controversies score, which measures the amount of ESG-based controversies a company has faced. The calculations based on the Fama and French (2015) five-factor model show that there is still potential for an investor to achieve a significant outperformance. Even though a value-weighted investing strategy does not show any significant over- or underperformance and therefore confirms many of the previous literature findings (see Halbritter and Dorfleitner, 2015), we can find some noteworthy results.

First of all, the inclusion of the controversies score in an ESG-based portfolio selection approach enables for a simple implementation as a way to quantify and evaluate the absence of a certain event, namely an ESG-based scandal, which might help to improve the information efficiency of the market with regard to the absence of these. Furthermore, from an investor's standpoint, having a "clean coat" with regards to controversies is especially profitable for smaller companies, as the absence of these scandals may be overlooked and incorrectly incorporated in the market prices. Thus, one might say that the respective companies "fly under the radar".

In addition, equally-weighted portfolio strategies based on worst TR and combined scores show significant outperformance, which leads to the conclusion that for the respective (small) companies there are indications in favor of the trade-off theory. Moreover, the results hold true for various robustness checks such as the variation of cutoff levels or the splitting of the period under review. Besides the two standard approaches in the context of portfolio formation, namely value- and equally-weighting, we discover new potential in the rank-weighted strategy for investors, which leads to improvements in terms of both, alpha and level of significance, within most of the investigated portfolios. For investors who attach great importance to ESG ratings, this represents an enormous opportunity to reward better scoring placements of companies and additionally to gain higher returns.

In light of these findings, it must, however, still be considered that there are hidden opportunities for investors that can be exploited in order to benefit from ESG-based ratings. The empirical results and arguments provided above prove that it is worth remaining vigilant concerning this issue.

Chapter 3

How socially irresponsible are socially responsible mutual funds? A persistence analysis

This research project is joint work with Gregor Dorfleitner (University of Regensburg) and Ralf Laschinger (University of Regensburg). The paper has been published as:
Dorfleitner, G., Kreuzer, C., Laschinger, R. (2021), How socially irresponsible are socially responsible mutual funds? A persistence analysis. *Finance Research Letters*, 43, 101990.

Abstract Based on a dataset of over 400 fund compositions in the years 2003–2018 this paper analyzes the persistence of controversies scores and environmental, social, governance (ESG) scores in socially responsible US mutual funds. As measurements for corporate social irresponsibility as well as corporate social responsibility activities, it is shown that US mutual funds exhibit controversies and ESG persistence in the short and longer-term. When examining the relationship between controversies and ESG scores in comparison with management fees, it becomes apparent that higher-paid managers achieve better results regarding controversies scoring but worse results regarding ESG scoring, compared to lower-paid managers.

Keywords ESG, controversies, socially responsible investing, corporate social responsibility

3.1 Introduction

For many years socially responsible investing (SRI) has been studied from a practical and academic point of view in a variety of ways. While many companies focus exclusively on promoting and maintaining their corporate social responsibility (CSR) activities, the question remains regarding to what extent corporate social irresponsibility (CSI) is avoided. Even if in today's society CSR activities of companies are gaining increasing attention, since being socially responsible entails not only doing a lot of "good" but also actively avoiding many more "bad" in terms of social irresponsible or unethical behavior, such as environmental scandals or business ethics controversies (see Lin-Hi and Müller, 2013). After all, an examination of CSR criteria alone is not sufficient to discover to which extent certain CSR activities are only used to make amends for past "sins", insure against possible subsequent CSI (see Kang et al., 2016), or remain part of sustainable corporate policy. Investors who value sustainability, but do not wish to make investment decisions themselves, often choose socially responsible (SR) mutual funds. These investors in particular put their trust in managers of SR mutual funds to act responsibly as well as to make forward-looking and sustainable investment decisions. Thus an additional evaluation concerning CSI criteria is not only interesting from an investor's point of view but should be of particular importance for SR mutual funds.

Dorfleitner et al. (2020) examine the relationship between corporate social performance (CSP) and corporate financial performane (CFP) regarding ESG (which stands for environmental, social and governance) rating as well as the *Thomson Reuters Controversies Score*[1] within the context of portfolio selection. By collecting and evaluating negative media stories from global media sources, this controversies score offers a new dimension with which to measure ESG-based scandals caused by the corporate behavior of the company under consideration.

In this study, we not only examine the ESG and controversies persistence of US mutual equity funds as a new dimension of ESG and a measurement for CSI, but also provide an assessment of this social performance (SP) for investors. This paper is the first to analyze and evaluate the persistence of US mutual equity funds regarding CSI as measured by an (ESG) controversies score based on an evaluation of the respective companies in their historical stock holdings.

Naturally, an important question with theoretical and practical consequences is whether SRI involves a direct performance benefit for investors. Concerning the academic literature, controversial results can be observed. While some authors show an outperformance of investors who apply socially responsible screens methods on their portfolios (see Kempf and Osthoff, 2007; Edmans, 2011), others ascertain neither performance benefits (Statman

[1]The scores are currently published by Refinitiv.

and Glushkov, 2009) nor negative performance (Hong and Kacperczyk, 2009).

Among SRI strategies some authors investigate the influence of positive or negative ESG based events and investors' reactions (Krüger, 2015) or the influence of bad media on firm value (Lundgren and Olsson, 2009).

In addition to the evaluation of individual companies based on ESG criteria, the importance of so-called sustainable funds is steadily increasing. According to the Report on US Sustainable, Responsible and Impact Investing Trends (see US SIF, 2018) the AUM of SRI assets is $12.0 trillion in the US market only.

Consequently, another major field of the SRI literature deals with the financial performance of SR mutual funds compared with conventional mutual funds. However, the majority of authors reports that no performance differences are achieved here (see, i.e., Statman, 2000; Bello, 2005; Bauer et al., 2005; Cortez et al., 2009; Utz and Wimmer, 2014). Furthermore various other authors investigate financial performance persistence of mutual funds (Hendricks et al., 1993; Carhart, 1997; Muñoz et al., 2013). In summary, hitherto no significant sustained overperformance of fund managers has been demonstrated in these studies either.

By analysing the ESG performance of mutual funds, Wimmer (2013) finds that the overall ESG rating based SP of SR mutual funds is persistent only for a short period and decreases after several years. Thus, in particular, ethically motivated value-driven investors must be attentive when selecting SR mutual funds to ensure that their requirements for sustainability are maintained in the medium and long-term. While there may be varying reasons for this group to invest in SR funds, Barreda-Tarrazona et al. (2011) show that investors who place a high value on sustainability also invest significantly more in SR funds once they are informed and convinced of their SR nature, even if they expect a lower return compared with a non-SR alternative.

This study is specifically aimed at investors who are primarily interested in responsible and sustainable investment criteria (Bollen, 2007; Renneboog et al., 2008) as opposed to performance-optimized portfolios. Since corporate social responsibility (CSR) criteria of companies are measured and quantified by ESG scores, we measure the ESG scoring of a SR mutual fund holding by weighting the ESG-scores of the individual assets. In the same way, we measure the controversies scoring of a SR mutual fund. To the best of our knowledge, we are the first to investigate the controversies score persistence of mutual funds by using a score based on ESG media scandals.

In this paper, we demonstrate that the short and long-term CSI and ESG persistence of SR mutual funds are preserved. Moreover, we show that the controversies-based SP of high-paid managers surpasses that of the lower-paid managers, whereas their ESG-based SP is clearly worse.

Finally, it becomes evident that funds with a very high ESG rating tend to have low controversies ratings and vice versa. From an ethical investor's point of view, it becomes apparent that they have to choose one side in their investment decisions.

Our work is organized as follows: In Section 2 we discuss our data set and introduce the methodology. The results are presented and discussed in Section 3.3. The Section 3.4 concludes this paper.

3.2 Data and methodology

To create our database we commence with a list of sustainable and responsible and impact mutual funds from the US SIF website, which provides a sample of all SRI classified mutual funds in the United States (https://charts.ussif.org/mfpc/). Since the current version of this website only covers open funds, we also include closed mutual funds, collected from former US SIF mutual funds lists, to prevent our results from survivorship bias. Next, we connect this list to the 'CRSP Survivor-Bias-Free US Mutual Fund Database', which contains various information, such as the exact holdings of the respective funds on previous reporting dates from 2003–2018. Furthermore, we add following two ESG based scores for the individual funds' compositions.

As the world's largest ESG rating database, we choose Thomson Reuters ESG score, which evaluates a company's environmental, social, and governance performance. It is calculated with the use of ten main themes (including resource use, innovation, emissions, human rights, workforce, management) based on publicly available company-reported data. These categories receive an individually measured category score and are weighted in the associated ESG pillar score. The aggregated pillar scores result in the overall ESG score, which is ranked by percentile and benchmarked against the industry (for more details regarding calculation see Refinitiv, 2021). In addition, we use the *Thomson Reuters Controversies Score*. As Thomson Reuters' latest scoring methodology, this score investigates firm controversies and adds a new dimension to previous approaches by capturing negative media stories from global media sources. In addition, it is benchmarked on respective industry groups (see Refinitiv, 2021).

Both scores range from zero to one hundred, and are interpreted in such a way that the higher the score the better the respective ESG or controversies rating. Table 3.1 shows the descriptive statistics of our data set.

To measure the overall ESG and controversies score of a fund on a particular reporting date, we weight the latest available scores before the reporting date concerning the individual securities weightings of the fund's composition. This calculation follows Dorfleitner et al. (2012) regarding the social return S_P of portfolio holding, which satisfies the portfolio

Table 3.1: Descriptive statistics.

Score	N	Mean	St. Dev.	Min	Max
ESG Score	422	53.5477	10.0938	27.2809	77.3793
Controversies Score	422	74.3090	13.2898	37.2108	100.0000

This table presents the mean, standard deviation, minimum, and maximum values of the TR, controversies scores of the full dataset.

additivity property

$$S_P = \sum_{i=1}^{N} x_i \cdot S_i,$$

where $\vec{x} = (x_1, \ldots, x_N)^T$ represents the portfolio weights of the assets with $\sum_{i=1}^{N} x_i = 1$.

Although both scores are available for an average of 2000 securities, we only take equity funds into account for which more than 60 percent[2] of the fund's holding is covered by both scores[3]. Thus we obtain a database of over 60 different funds and over 400 fund compositions[4] with ESG and controversies score coverage. The number of fund compositions per year is displayed in Table 3.2.

Table 3.2: Number of fund compositions per year.

Year	2003	2004	2005	2006	2007	2008	2009	2010
N	10	12	14	21	35	28	27	25
Year	2011	2012	2013	2014	2015	2016	2017	2018
N	25	27	28	26	37	36	36	35

This table reports the number of fund compositions per year of our dataset.

The possible problem of overlapping stock holdings of individual funds is no issue in our sample, as can be seen from the wide variety of observed portfolio scores.

Following a similar approach to Wimmer (2013), we categorize the funds concerning their overall ESG and controversies score on a yearly basis. For this purpose we construct four equally weighted portfolios for each score, one for each quartile of the funds' ratings on a yearly basis. This yields to eight portfolios, four built on ESG scores and another four built on controversies scores. Here, ESG portfolio 1 contains the quarter of all funds with

[2] As a robustness test, we examine various percentages (50%, 70%). The results remain materially unchanged.

[3] As another robustness test, we calculated our portfolios with all available ESG or controversies ratings. Again, the results remain materially unchanged.

[4] The number of available fund compositions increases from about 10 in earlier years to about 35 in later years. Over 83% of the funds exhibit at least 3 consecutive observations. On average, our data base covers 6 observations per fund.

the lowest ESG score, while controversies portfolio 1 contains the 25 percent of all funds with the lowest controversies score. Analogously ESG/controversies portfolio 4 contains the 25 percent of all funds with the respective best overall score.

3.3 Results and discussion

The presented results are sorted according to their time horizon, which reflects possible investment horizons. Following Wimmer (2013), we define short-term as being one year, mid-term horizon as being one to three years and long-term as three years or more.

3.3.1 Short-term persistence

To examine the short-term persistence of mutual funds, we examine a contingency table of current and subsequent one-year ranking transition. Table 3.3 presents the probability of a fund in a specific ESG or controversies rank portfolio of falling in each rank portfolio in the subsequent year. The data shows that for a one-year persistence, the funds' ESG and controversies score remains largely unchanged, especially for portfolios 1 and 4 of each score. Thus the highest probabilities of remaining in the top quartile are observed for funds in the top quartile (approximately 79% for ESG and 73% controversies portfolios) as well as for funds in the bottom quartile of remaining in the bottom quartile (approximately 85% for ESG and 73% for controversies portfolios).

Table 3.3: Contingency table of controversies and ESG portfolios.

Score		1	2	3	4
Controversies	1	0.7333	0.2111	0.0555	0.0000
	2	0.2841	0.4204	0.2841	0.0114
	3	0.0919	0.3563	0.4023	0.1494
	4	0.0109	0.0652	0.1956	0.7283
ESG	1	0.8469	0.1122	0.0408	0.0000
	2	0.0330	0.6374	0.2637	0.0659
	3	0.0000	0.1363	0.5454	0.3182
	4	0.0125	0.0125	0.1825	0.7875

This table displays the contingency table of initial and subsequent fund controversies and ESG quartile rank rating. In every year between 2003 and 2018, the observed funds are ranked in one of the four rank portfolios. These rankings are connected to the subsequent fund quartile ranking.

All things considered, additionally to the findings of Wimmer (2013) for the ESG rating of mutual funds, we have an indication that "winners stay winners" and "losers stay losers" also applies to the controversies score.

Considering the Spearman correlation test for ESG and controversies scores of the portfolios we measure a significant non-zero correlation between the original and subsequent ranking.

3.3.2 Mid- and long-term persistence

From the perspective of an investor it is of course interesting to find out how far the ESG and controversies scoring of SR mutual funds persist throughout the following years. For this purpose, we calculate the average scores of the eight portfolios in their initial year as well as for the subsequent four years, whilst refraining from any rebalancing. Again, portfolio 4 contains the top quartile funds with the highest ESG and controversies scores and portfolio 1 contains the bottom quartile funds with the respective lowest ratings.

Figures 3.1 and 3.2 show the descriptive data for the development of the four ESG and controversies portfolios in the subsequent four years. When considering the controversies portfolios it becomes evident that the top portfolio remains by far the best portfolio in the following years. The other three portfolios show no major changes in their score developments and thus maintain their ranks.

By regrading the ESG portfolios, it can be seen that the portfolios 2 to 4 converge with regard to their overall ESG scores. Even with a split into two sub-periods (2003–2010, 2011–2018), the effect described above remains intact. However, in contrast to Wimmer (2013) we cannot find a change in the rank formation of the ESG portfolios when considering the Thomson Reuters ESG score.

Again, when considering the Spearman correlation test for ESG and controversies rank portfolios, we measure a significant non-zero correlation between the original and each of the subsequent four rankings.

3.3.3 ESG and controversies scores vs. Management Fee

Chevalier and Ellison (1999) show that the performances of mutual funds are related to the characteristic of fund managers, such as behavioral differences, age, and education. However, especially for mutual funds, not only the financial performance is important, but also the quality of the investment decisions with regard to various sustainability criteria. In particular, it is value-driven ethical investors who demand the highest possible standards, which poses a behavioral challenge for managers, especially since financial performance must not be completely neglected, in order to remain competitive. For instance, Dorfleitner et al. (2020) show that particularly small companies with low ESG ratings achieve significant overperformance. Thus it remains questionable whether managers avoid investment opportunities despite an attractive performance in order to achieve a high ESG

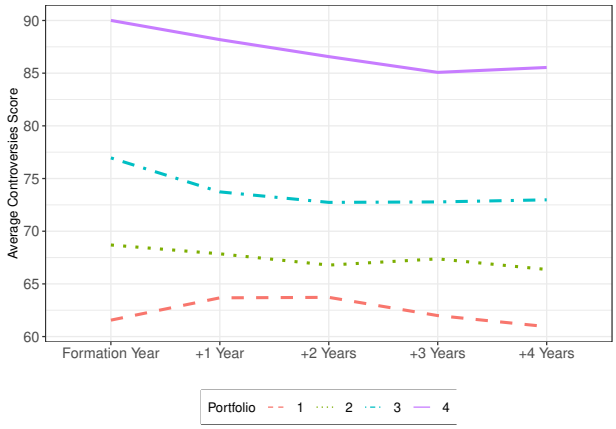

Figure 3.1: Long-term persistence of controversies portfolios.

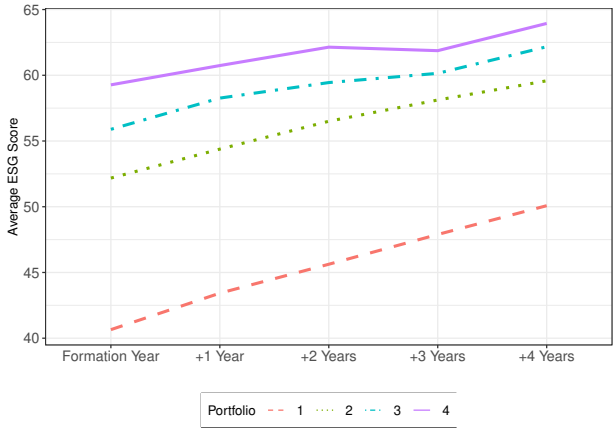

Figure 3.2: Long-term persistence of ESG portfolios.

Description: For the years from 2003–2018 we rank each fund into equal-weighted quartile portfolios based on their overall ESG scores as well as in equal-weighted quartile portfolios based on their overall controversies scores. The lines represent the ESG and controversies score of the four rank portfolios in their formation years and the subsequent four years without any changes to their formation. Funds that initially achieve the highest ESG or controversies ratings are contained in portfolio 4 and those with the lowest ratings appear in portfolio 1.

standard.

To investigate whether high-paid managers of SR mutual funds demonstrate better ESG or controversies social performances of their funds than the lower priced ones, we add the respective fund management fee[5] of over 300 different funds to our dataset. The descriptive statistics of the additional dataset are shown in Table 3.4.

Again, we divide our portfolios into eight rank portfolios (four for each score). Portfolio 1 includes the funds with the 25% low priced managers whereas portfolio 4 covers the 25% of high-paid managers. Afterwards we calculate the overall ESG and controversies scores of the respective portfolios. Note that due to the data limitation, sufficient observations

[5]Note that the management fee does not include the expense ratio.

Table 3.4: Descriptive statistics.

	N	Mean	St. Dev.	Min	Max
Management Fee	306	0.5675	0.2062	0.0500	1.0000

This table presents the mean, standard deviation, minimum, and maximum value (in %) of the management fee dataset.

(minimum two for each rank portfolio) are only available as from 2007. Similar to the procedure above, we examine the initial scores as well as the development over the ensuing years.

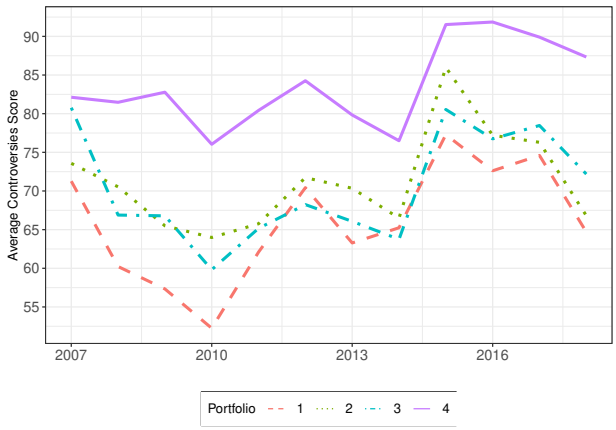

Figure 3.3: Controversies score of fee portfolios.

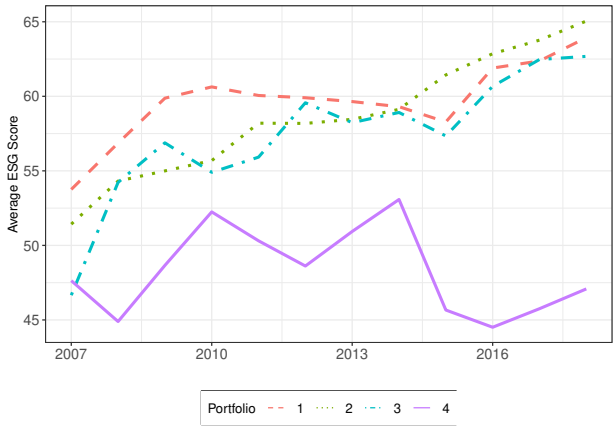

Figure 3.4: ESG score of fee portfolios.

Description: In the years from 2003–2018 we rank each fund in equal-weighted quartile portfolios based on their management fee. The lines represent the ESG and controversies score of the four rank portfolios from 2003–2018. Funds that are managed by the highest-paid managers are contained in portfolio 4 and those which are managed by the lowest-paid managers are contained in portfolio 1.

By considering Figures 3.3 and 3.4 we can see some noteworthy results: On the one hand, the highest paid managers show the best ongoing controversies scoring, whereas the scores of the other three portfolios from 2014 onwards no longer exhibit any major differences.

On the other hand, this effect changes dramatically when considering the ESG scoring of the rank portfolios. Here, the highest paid managers show by far the lowest ESG score. Surprisingly, in the majority of years, the highest overall ESG scoring can be detected by the lowest paid managers. After 2015, there are again no major differences within the rank portfolios 1 to 3.

Note that some of the results in early years are driven by few funds per portfolio due to the limited sample size.

3.3.4 ESG score vs. controversies score

Another interesting question for investors is whether SR mutual funds with high ESG ratings also achieve high controversies scores and vice versa.

For this purpose we begin with the four ESG rank portfolios and calculate the respective controversies ranking and vice versa. Again we examine the developement of the subsequent four years. Concerning Figures 3.5 and 3.6, we surprisingly detect a major change in the ranks of the respective portfolios. Conversely, the best ESG rank portfolio exhibits one of the worst controversies scores and, on the other hand, the worst ESG rank portfolio shows by far the best ongoing controversies scores. A similar picture can be seen when considering the ESG scoring of the controversies rank portfolios. Once more, the best controversies rank portfolio shows the worst ESG scoring. One reason for this development is the generally rather negative correlation of ESG and controversies score (see Dorfleitner et al., 2020). It is therefore difficult for fund managers to be leaders in both ratings. Another possible explanation could be that they maintain their investment policies concerning the focus on either ESG or controversies ratings. We leave a clarification of this matter to further research.

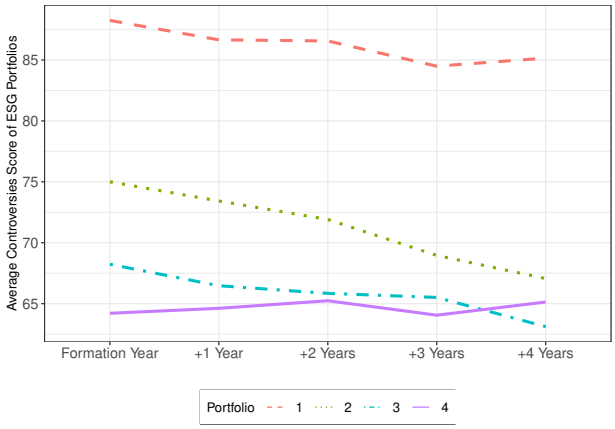

Figure 3.5: Controversies Scoring of ESG rank portfolios.

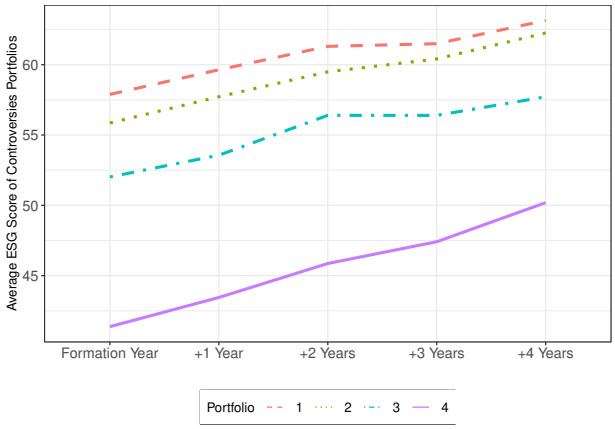

Figure 3.6: ESG Scoring of controversies rank portfolios.

Description: In the years from 2003–2018 we rank each fund in equal-weighted quartile portfolios based on their overall ESG scores as well as in equal-weighted quartile portfolios based on their overall controversies scores. The lines represent the ESG or controversies score of the four controversies, respectively ESG rank portfolios in their formation year and the subsequent four years without any changes to their formation. Funds that initially achieve the highest ESG or controversies ratings are contained in portfolio 4 and those with the lowest ratings are contained in portfolio 1.

3.4 Discussion and conclusion

First, the results shown above can be considered as good news for value-driven investors. With the growing popularity of ESG ratings, SR funds are obviously also pursuing clear and relatively consistent investment policies. From the data it becomes apparent that this effect can be seen for both ESG and controversies portfolios, not only in the short term but also in the medium and long term. This is particularly useful for investors who do not wish to actively rebalance their investments.

Second, the question arises as to why particularly high-paid fund managers perform well in controversies ratings but worse in the ESG rating of their funds than lower-paid managers.

On the one hand, media attention and influence is more present than ESG reporting and ratings. Thus, particularly high-paid managers may try to avoid negative reporting upon the companies in which they invest. On the other hand, this may also prevent sudden performance losses due to media scandals. Therefore, high-paid managers try to avoid companies that are prone to scandals by placing special emphasis on companies that exhibit good controversies SP compared with the remainder of their respective industry groups. Nevertheless, a certain level of profitability of the investments needs to be reconciled with sustainability considerations.

Third, it becomes apparent that there are difficulties for SR funds to become leaders in both ESG and controversies. These results are not surprising and remind us of the "the higher you fly the harder you fall" hypothesis (Dorfleitner et al., 2020). Especially here, investors must decide to what extent one or the other rating is more important to them and to find a personal preference.

All in all, our results show that there are certain similarities but also differences between the ESG and the controversies rankings of mutual funds. The results can be summarized and delimited as follows:

In the short-term development almost no differences can be seen when comparing the initial and subsequent portfolio rankings of both controversies and ESG scoring. Also, in the longer-term (Section 3.3.2), many ESG portfolios are virtually identical or almost converge in their ESG scores. The respective controversies portfolios show a rather delimited and constant development, especially in the top quartile, but the rankings also remain unchanged.

When we examine the relationship between ESG and controversies scores and management fees (Section 3.3.3), we find that higher-paid managers achieve better results regarding controversies scorings but worse results regarding ESG scoring. Again, the effect is particularly evident in the top quartile, whereas the other quartiles converge over time.

Last but not least, we find evidence of the fact that the controversies and ESG scores of mutual funds show clearly opposing developments. Funds with high ESG ratings tend to have comparatively low controversies ratings and vice versa.

Despite various robustness tests, the results of this first study are somewhat limited due to the limited sample size. Note, that in particular in the early years (2003–2005), some of the rank portfolios consist only of few funds. Therefore this may affect some of the early stage observations, but this effect disappears after only a few years. This work is intended to provide a first step toward examining CSI of SR equity funds and naturally leaves some space for further research based on a larger data basis.

If a value-driven investor considers the controversies component in addition to ESG criteria, certain investment decisions become more difficult.

Scandals, in particular, are difficult to predict, occur across all industries, and may have an enormous financial impact in the short- or even mid-term (see, e.g., emission scandal, oil spill scandal, accounting scandal).

Ratings such as controversies scores help investors to assess companies regarding their vulnerability to controversies relative to companies in the same industry. Especially investors who value both ESG and controversies find themselves to be in some degree of conflict. Our work provides the first step towards a new aspect of ESG assessment, leaving space for further investigative research.

Acknowledgements

We would like to thank Dr. Maximilian Wimmer for useful discussions and support.

Chapter 4

It's not only size that matters: On the influence of policy, society, culture, and firm characteristics on corporate controversies

This research project is joint work with Gregor Dorfleitner (University of Regensburg) and Christian Sparrer (University of Regensburg).

Abstract Using industry-based controversy ratings for 5,700 companies from 44 countries, and a hybrid panel data model which allows us to separate the within-firm and the between-entity effects, our empirical approach examines a broad spectrum of diverse political, cultural, societal, and firm-specific variables that influence patterns of unethical corporate behavior and their disclosure. We argue that companies tend to have fewer controversies if there is a high level of institutional pressure or if corporate controversies pose a high-level threat to organizational legitimacy. Moreover, in highly moral societies that closely monitor corporate behavior, more corporate controversies can be observed. Our results are further confirmed through the use of an alternative model which examines the number of scandals mentioned in the media in contrast to the controversy rating.

Keywords Controversy, hybrid model, institutional theory, legitimacy theory, corporate scandal

4.1 Introduction

The departure from Friedman's (1970) merely profit-maximizing shareholder-value orientation toward a stakeholder welfare concept that accounts for the diverse needs of employees, customers, and the broader society has become a core concept in the business world, in international politics, and also in academic literature. Corporate controversies such as environmental scandals, the exploitation of the workforce, or shareholder rights infringements, however, pose a direct threat to this concept of stakeholders' welfare orientation. This empirical study identifies different drivers of corporate controversies or, henceforth used synonymously, corporate scandals[1]. Based on our theoretical framework, we show that various nation-level and company-specific variables affect unethical corporate behavior and the disclosure thereof. These insights are important not only for value-oriented investors to identify ethical firms but also for the decision-process of managers. They may be apt to guide regulators and policy-makers and to sensitize general societal awareness towards (un-)ethical behavior.

The academic literature on corporate social responsibility (CSR)[2] has seen a tremendous growth over recent years. This focus on CSR-related issues is, however, biased toward the notion of *good* social performance (Aouadi and Marsat, 2018), while mostly neglecting the occurrence of CSR-related scandals. Corporate controversies not only harm a firm and its shareholders directly, as the disclosure of controversies usually comes with a stock price decline. These events can also inflict severe damage to the environment, as was the case with the Deepwater Horizon oil spill in 2010, or to society, which can be illustrated by the loss of public trust in regulatory agencies after they failed to detect Wirecard's accounting fraud in 2020.

This paper extends the understanding of irresponsible corporate behavior by shifting the academic focus toward publicly perceived, CSR-based corporate controversies. We are, therefore, the first to answer the following key question: Which nation-level institutions or determinants and which company-specific factors support patterns of unethical corporate behavior or foster the disclosure of corporate social irresponsibility? Drawing upon institutional theory, we argue that high institutional and stakeholder pressure can alter a firm's behavior to adopt ethical standards. By employing the legitimacy theory, we hypothesize that, if corporate controversies pose a direct threat towards organizational legitimacy, companies will act more ethically. Furthermore, we argue that societies with high moral standards that closely monitor and scrutinize corporate behavior are more

[1] In general, a scandal describes a clearly deplorable behavior, while the term controversy comprises two possible opposite perspectives. Nevertheless, in the context of companies, every controversially discussed issue harms the reputation and can therefore be considered a scandal.

[2] Following Liang and Renneboog (2017), we define CSR as business activities that enhance social welfare, regardless of accordance with profit maximization.

likely to disclose corporate controversies. Despite its rich theoretical foundation, this paper refrains from formulating hypotheses and is intended as a first empirical study to identify a broad range of different drivers of corporate controversy.

This paper proceeds as follows. First, we lay the theoretical foundation for our research and describe the relationship between different nation- or company-level dimensions and corporate controversies. Second, we describe our data and empirical models. Third, we present and discuss the empirical results and apply several robustness checks to verify our results. Finally, we conclude and address the aspect of potential further research.

4.2 Theoretical development

4.2.1 General framework

To establish our multi-level framework, a formal definition of the term *corporate controversy* is indispensable. In order for a corporate controversy to occur, we follow the definition of a media scandal of Thompson (2005) and posit two aspects: (1) the *unethical corporate behavior* itself and (2) the process of *societal disclosure of corporate controversies*. This implies that stakeholders must perceive this pattern of unethical corporate behavior and it has to cause a certain level of indignation in order to be considered relevant enough for them to act (Aouadi and Marsat, 2018; Weick et al., 2005). In some cases, such as the Volkswagen emission scandal (commonly known as "Dieselgate") around 2015, the actual unethical corporate behavior remains hidden for a long time before the corporate misconduct is publicly disclosed.

We define unethical corporate behavior as corporate environmental, social, or governance offenses[3] (Aouadi and Marsat, 2018) that are caused by a company either willingly and knowingly or because of negligence. Examples of unethical corporate behavior comprise offenses such as Nestlé's child labor controversy in 2005, the Deepwater Horizon oil spill in 2010, and Siemens AG's bribery of the Argentine government in 2008. The second requirement, i.e., the process of societal disclosure, includes the perception, disapproval, and publication of this unethical behavior. As political, societal, and cultural institutions directly influence values and ethical standards (Alas, 2006; Beekun and Westerman, 2012), they also shape the perception and definition of unethical corporate behavior. In some countries, patterns of unethical corporate behavior are neglected and may thus remain concealed, whereas, in others, the same kind of bad behavior is uncovered, condemned, and publicly disseminated.

Figure 4.1 depicts our theoretical model for the occurrence of corporate controversies.

[3]We align ourselves with Godfrey (2005) who also use the term *offense* instead of *crime* due to the fact that a scandal is not always the result of an illegal action.

Following recent literature (see, e.g., Baldini et al., 2018; Ioannou and Serafeim, 2012), we separate our variables into nation-level institutional and company-specific factors. Thus, we examine the influence of country- and company-specific determinants on both the unethical behavior and the process of disclosing scandals.

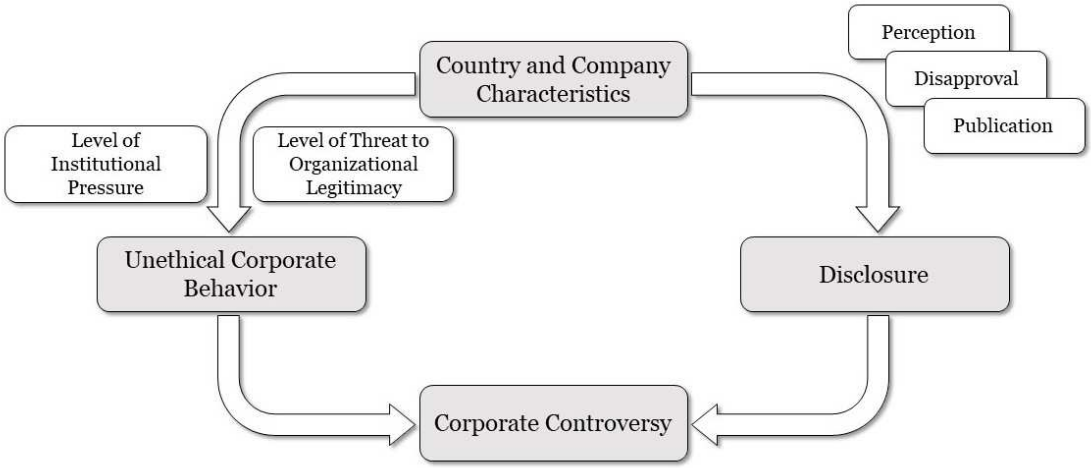

Figure 4.1: A theoretical model of corporate controversies.

Regarding the channel of unethical behavior, we mainly draw upon institutional theory (see, e.g., DiMaggio and Powell, 1983; Meyer and Rowan, 1977; Oliver, 1991) and legitimacy theory (see, e.g., Aldrich and Fiol, 1994; Ashforth and Gibbs, 1990; Suchman, 1995) to explain cross-national variations and intra-industrial differences. We argue that a high level of institutional pressure forces companies to adopt ethical behavior and that the threat of losing legitimacy through involvement in a corporate controversy incentivizes firms to behave ethically[4].

According to institutional theory, organizational choice is limited through the application of pressure of external institutions on an organization[5]. Organizations must respond to institutional expectations and demands in order to survive (DiMaggio and Powell, 1983; Meyer and Rowan, 1977; Pfeffer and Salancik, 1978). This strategic response takes place even in the absence of an increase in internal organizational efficiency (DiMaggio and Powell, 1983), which clearly emphasizes that the position of power is ascribed to the institutions. Thus, institutional theory argues that the predominant motivation for ethical corporate behavior in general is mainly to address and respond properly to the tremendous

[4]Although some researchers (see, e.g., Baldini et al., 2018; Drempetic et al., 2019; Schaltegger and Hörisch, 2017) argue that institutional theory and legitimacy theory tend to overlap, our study focuses on the use of both theories separately, as this allows a more in-depth differentiation of potential drivers of unethical corporate behavior.

[5]Institutional environments not only comprise regulatory or cultural structures, governmental agencies, and laws (Oliver, 1991; Scott, 1987), but also stakeholders or interest groups (Drempetic et al., 2019) with both their behavioral norms and values (North, 1990).

pressure placed on the organization by customers, employees, shareholders, NGOs, or the government.

Legitimacy theory, on the other hand, primarily emphasizes the benefit of conformity to the needs and expectations of various stakeholders. Organizations that adhere to external rules and norms gain or retain legitimacy (Ashforth and Gibbs, 1990; Dowling and Pfeffer, 1975; Oliver, 1991), which Suchman (1995) defines as the assumption or generalized perception that an entity acts in a desirable, proper, or appropriate way and that these actions are in line with the societal system of norms, values, and beliefs. Society awards organizations with a "license to operate" as long as they fulfill societal needs and enhance societal welfare (Baldini et al., 2018). A lack of organizational legitimacy leaves the organization vulnerable to greater external scrutiny (Ashforth and Mael, 1989) or claims that the organization is unnecessary (Meyer and Rowan, 1977; Suchman, 1995). This can ultimately result in a revocation of the "license to operate" and pose a direct threat to the economic survival[6]. Particularly in the context of corporate controversies, the revocation of legitimacy manifests itself in various forms: socially responsible investors shun investment in unethical firms, value-oriented customers refuse to buy the firm's products or even call for a boycott, suppliers end contracts, or capital markets impede the access to financial resources.

Regarding the second channel of our model, i.e., the disclosure of controversies, societies with, for example, high moral standards will monitor corporate behavior more closely and uncover corporate controversies more easily. In societies with lower moral standards, however, perceived unethical behavior does not cause indignation and, thus, the scandal remains hidden. Finally, the combination of the actual unethical corporate behavior and its disclosure results in a perceptible corporate controversy.

4.2.2 Nation-level institutional and economic determinants

Many diverse political, societal, and economic stakeholders can influence a firm's behavior (Campbell, 2007; Ioannou and Serafeim, 2012) by shaping the formation of societal norms and values (Alas, 2006; Beekun and Westerman, 2012). Inspired by Whitley's (1999) well-established "national business systems" (NBS) institutional framework, we capture these effects within a set of nation-level determinants which comprise four dimensions: (1) the political and legal system, (2) the cultural system, (3) the labor and education system, as well as (4) the financial system. Variables within these dimensions may vary over time and between different countries, but not between different entities located in the same country.

[6]Organizational behavior may diverge temporarily from external expectations and rules or norms, but can simultaneously maintain legitimacy, because the divergence either remains hidden from societal perception or incites no public disapproval. This, once again, emphasizes the importance of interweaving unethical corporate behavior and its disclosure.

Political and legal system. It is indispensable to address corporate controversies in the context of a country's political and legislative environment which mainly covers formal institutions (North, 1990), such as law enforcement and valid regulations. These formal institutions play a key role in the development of the relationships of companies with their key stakeholders (Ioannou and Serafeim, 2012; Liang and Renneboog, 2017). Whitley (1999) and Rodriguez et al. (2005) stress the importance of the government, which regulates market entries and exits and constrains the activities of economic actors.

As posited by institutional theory, strong and stable national governments can pressurize firms into adopting proper ethical behavior through enacting and enforcing new laws (Aguilera et al., 2007). These laws define clear societal expectations concerning ethical corporate behavior, which are then reinforced by consumers, employees, NGOs, and other actors (Aguilera et al., 2007; Kagan et al., 2003). Besides the enactment of laws, a strong level of law enforcement and effective penalization for noncompliance are key components to ensure ethical corporate behavior. Companies only adopt reasonable social behavior when they fear detection of and penalization for noncompliance (Becker, 1968; Coluccia et al., 2018; Kagan et al., 2003; Oliver, 1991). In the long run, once appropriate laws have been passed, and their enforcement has been ensured, firms will be more likely to comply with institutional pressure toward social responsibility (Oliver, 1991).

The power of the government to push companies toward ethical standards depends not just on the presence or the amount of regulation per se (Campbell, 2007), but also on the actual, or at least perceived, societal conformity to these standards. Suchman (1995) states that legitimacy depends on the societal definition of desirable actions. In countries with high levels of corruption or in countries with an unstable political system, dubious actions, such as bribery, often appear to be appropriate forms of behavior. They award organizations legitimacy (Suchman, 1995) and firms may engage in unethical practices in order to achieve a competitive advantage (such as cutting costs at the expense of product safety). This ultimately forces other companies to adopt these unethical standards to remain competitive (Ioannou and Serafeim, 2012), effectively creating a culture of acceptance (Argandoña, 2001), in which corporate misconduct does not pose a threat to organizational legitimacy. Turning to the societal disclosure dimension, unethical behavior is not perceived as a controversy in these countries but instead urges other companies to emulate this behavior to ensure economic survival (Rodriguez et al., 2005). The societal inability or the unwillingness to disclose corporate misconduct is even further strengthened if the government fails to ensure and encourage both free speech and political participation.

Cultural system. Informal institutions (North, 1990), such as a country's national culture, structure the communication and relationship between business partners, employees, and the firm itself (Whitley, 1999), directly influence the values and moral beliefs an individual possesses (Alas, 2006), and strongly impact the perception of ethical standards and

actions (Beekun and Westerman, 2012). These moral standards are not established by a superordinate political body, but instead obtain their validity through inner justification (Velasquez and Velazquez, 2002).

From an economic perspective, institutional theory implies that organizations either respond to an increase in cultural pressure through the adoption of these cultural values (Dickson et al., 2004; Waldman et al., 2006a), or by changing their behavior (Oliver, 1991). For example, Crossland and Hambrick (2011) and Ioannou and Serafeim (2012) argue that a nation's culture determines management discretion and decision-making processes.

Following Hofstede's cultural framework (see Hofstede, 2001; Hofstede et al., 2010), some cultural traits, such as the perception and acceptance of unequally distributed power and the authority of leaders within society (i.e., the level of power distance) or the tolerance of unknown and ambiguous situations, create a cultural environment, in which managers and decision-makers must comply with institutional pressure and act in accordance with social norms in order to retain their position of power and legitimacy (Ioannou and Serafeim, 2012; Waldman et al., 2006a; Waldman et al., 2006b; DiMaggio and Powell, 1983). Other cultural characteristics, such as the degree of individualism, or the orientation towards the societal values of assertiveness, toughness, and success, encourage managers to pursue personal goals rather than the needs of society and long-term relationships with stakeholders (Ringov and Zollo, 2007; Waldman et al., 2006a; Vitell et al., 1993; Rallapalli et al., 1994). So they feel less pressure to act ethically.

Again turning to the disclosure of controversies, cultural norms also affect the perception of undesired and unethical behavior (Beekun and Westerman, 2012) and consequently the disclosure of corporate controversies. In collectivist societies, or in societies with high levels of power distance, criticism of and the discourse with authorities is undesired (Crossland and Hambrick, 2011; Tyler et al., 2000). Loyalty towards the actions and beliefs of the personal peer-group is mandatory (Hofstede and McCrae, 2004). Individuals as well as stakeholders accept the decisions of their superiors without questioning them (Hofstede et al., 2010; Williams and Zinkin, 2008), even if one disagrees with them. As a result, controversies are less likely to be publicly disclosed.

Labor and education system. Besides political and cultural institutions, Whitley (1999) also emphasizes the importance of labor and education systems. Good access to trained and motivated workers is crucial for the survival of any firm. Especially if this availability is limited in a certain country, companies have to compete for highly qualified workers.

Prior studies (e.g., Brammer et al., 2007; Greening and Turban, 2000; Ioannou and Serafeim, 2012; McWilliams and Siegel, 2001; Peterson, 2004; Sánchez, 2000; Turban and Greening, 1997) argue that firms associated with high levels of CSR attract and retain higher-qualified workers. Highly qualified employees expect a certain level of ethical be-

havior from the company (e.g., labor safety, non-discrimination policies in the recruitment process) because they would have a more positive self-concept. Consequently, highly skilled prospective job applicants prefer to work for ethically oriented firms (Brammer et al., 2007; Greening and Turban, 2000; Mueller et al., 2012; Peterson, 2004). In countries with limited availability of skilled labor, however, corporate controversies pose a high threat to organizational legitimacy, as they reduce the attractiveness of potential employers.

As job choice is a process with imperfect information (Bauer and Aiman-Smith, 1996; Rynes et al., 1991), firms can use different channels, such as their CSR policies, to communicate their compliance with ethical values. This increases the desire of potential job applicants to pursue the job (Greening and Turban, 2000; Turban and Greening, 1997). However, in this context, corporate controversies send negative signals concerning the company and its values, which can scare potential high-qualified job applicants off (Bauer and Aiman-Smith, 1996; Rynes et al., 1991). This, in turn, pressures firms into behaving ethically correctly in order to avoid disadvantages related to the recruitment process. Finally, we do not expect a direct effect of the labor market on the channel of societal disclosure.

Financial and economic system. Last, financial and economic systems determine how capital is made available to firms and how it is priced (Whitley, 1999), which is of vital importance to any firm. While we address the role of financial markets and their interactions with firms on the firm-level later, we only consider general economic conditions on the country level, as they also shape corporate behavior, such as market competition or the state of economic development.

The level of competition can directly affect corporate conduct, as firms alter their behavior to achieve advantages in the competition for customers, investors, or capital lenders. Campbell (2007) and Shleifer (2004) argue that companies operating within tight competitive environments may abandon ethical standards (such as product safety or fair payments) to cut costs and to guarantee economic survival. As opposed to that, firms that do not face any competitive pressure may have little incentives to act ethically, because a firm's reputation or customer loyalty will not likely affect sales or profitability and, thus, corporate controversies do not pose a threat to organizational legitimacy (Campbell, 2007).

Moreover, the general state of economic development and, consequently, the standards of living and education determine the perception of ethical standards as well as the disclosure of potential controversies. While their basic needs appear largely to be satisfied, people in developed countries tend to pursue additional higher goals, such as sustainable management of natural resources or no sexual or racial discrimination (Liang and Renneboog, 2017; Ho et al., 2012). According to institutional theory, the pressure on companies to comply with these societal needs and to behave ethically strongly increases the further a country is economically and socially developed (McWilliams and Siegel, 2001; Muhammad et al., 2011). Regarding the disclosure dimension of controversies, even a slight misconduct

by a company can cause dissatisfaction which will be publicly exposed and punished by the broader society.

The acceptable standards for corporate behavior may also spill over to other countries (Liang and Renneboog, 2017), and firms operating in more globalized countries must bow to the international pressure and adopt appropriate ethical standards to retain legitimacy.

4.2.3 Company-specific factors

In addition to country-related factors, we examine four micro-level dimensions that describe the relationship of a company with its key stakeholders: (1) CSR reputation, (2) firm size and visibility, (3) dependency on capital markets, and (4) the risk and return of a company. These company-specific drivers vary both over time and between different entities located in the same country.

CSR reputation. A high level of corporate commitment towards CSR establishes an ethical reputation and strategic value (McWilliams and Siegel, 2001; Porter and Kramer, 2002). Intuitively, this reputation is directly linked to corporate controversies.

Prior work (see, e.g., Bhattacharya and Sen, 2004; Gardberg and Fombrun, 2006; Godfrey, 2005; Godfrey et al., 2009; Knox and Maklan, 2004; Peloza, 2006) argues that companies with a good CSR reputation face alleviated sanctions following a controversy, which is mainly due to the positive effect of CSR and corporate citizenship building up relationship-based intangible assets[7]. For example, Fombrun et al. (2000) assert that CSR reinforces the network with various stakeholders (i.e., the government, customers, and employees) and, in case of a controversy, the involved corporation will receive support and backup from this network. Reinforced through the network created by CSR, stakeholders are much more inclined to "give the company a second chance". As a consequence, philanthropic companies face attenuated institutional pressure, which eventually reduces the threat posed by corporate controversies to their organizational legitimacy.

Besides these effects on unethical corporate behavior, the CSR reputation may also play a crucial role in the context of societal disclosure. Companies associated with a good CSR reputation draw more public attention towards their actions. Uncovering a major CSR-related scandal of one of the high-level CSR companies seems much more newsworthy, which is why external organizations such as NGOs or the media scrutinize them more closely.

[7]This type of relational wealth (see Business Ethics Quarterly, 2002) comprises aspects such as employees' affective commitment toward the company (Brammer et al., 2007; Fombrun et al., 2000; Mayer and Schoorman, 1992; Meyer and Allen, 1991; Peterson, 2004), societal and political legitimacy (Fombrun et al., 2000; Handelman and Arnold, 1999; Sánchez, 2000), competitive advantages such as supplier and customer loyalty and trust (Bhattacharya and Sen, 2004; Brown and Dacin, 1997; Fombrun et al., 2000), or brand equity (Gardberg and Fombrun, 2006; McWilliams and Siegel, 2000).

Firm size and visibility. As larger firms tend to be more visible among the broader society, they attract a wider spectrum of diverse stakeholders that influence the company (Fiss and Zajac, 2006; Hackston and Milne, 1996; Knox et al., 2005). Institutional theory suggests that large companies in particular are often confronted with diverse or even conflicting institutional demands (Ashforth and Gibbs, 1990; Oliver, 1991).

Based on DiMaggio and Powell (1983), who note that institutional pressure is a function of an organization's dependency on its institutional environment or its resources, larger companies, which are less reliant on a single stakeholder, can withstand external pressure more easily. Drawing upon legitimacy theory, both Meznar and Nigh (1995), and Pfeffer and Salancik (1978), emphasize that because the societal costs of revoking legitimacy from larger firms are far higher than those for smaller firms, larger companies are more resistant to the threat corporate controversies pose to organizational legitimacy. This may encourage larger firms to neglect or defy certain stakeholder demands and ultimately result in a conflicting or even unethical action that can lead to a corporate controversy.

Besides the actual unethical corporate behavior, the perception and disclosure of this corporate misconduct rely on third-party institutions (i.e., Media or NGOs) to publicly disseminate this information. Drempetic et al. (2019) and Schreck and Raithel (2018) state that firm visibility directly affects the amount of third-party information. Compared to less visible companies, high-attention firms are more newsworthy and are hence deemed to be more greatly exposed to public scrutiny (Aouadi and Marsat, 2018; Meznar and Nigh, 1995; Reverte, 2009; Salancik and Pfeffer, 1978; Servaes and Tamayo, 2013; Watts and Zimmerman, 1986). As an increasing number of stakeholders monitor the behavior of high-attention firms more closely, patterns of unethical behavior can be uncovered more easily.

Dependency on capital markets. Companies with financial resource shortages or firms with close ties to capital markets (i.e., firms with high levels of long-term debt obligations) exhibit a high dependency on capital markets. This dependency affects the propensity for corporate misconduct as well as its disclosure in many ways.

First of all, a high level of dependency on capital markets constrains unethical behavior, because corporate controversies can limit or impede the much needed access to finance. Consistent with institutional theory, if capital markets incorporate standards for ethical business behavior into their capital allocation process, these markets can hamper financing for socially irresponsible firms and eventually pressure them into adopting ethical standards. In this context, legitimacy theory suggests that a decrease in organizational legitimacy manifests itself in the form of impaired access to financial resources and a higher cost of capital, which is especially threatening for capital-constrained firms[8]. In contrast,

[8]Some authors ascribe this increase in cost of capital to a decreasing investor base for socially irresponsible firms (see, e.g., El Ghoul et al., 2011; Goss and Roberts, 2011; Heinkel et al., 2001; Hong and

firms with abundant slack financial resources are less dependent on capital markets, hampering the ability of capital markets to exert pressure on these companies. According to the free cash flow hypothesis (Jensen, 1986) and empire building theory, high levels of slack financial resources may cause agency problems (Servaes and Tamayo, 2014) and entice managers to pursue private benefits instead of a sustainable and ethical corporate behavior.

Second, Schreck and Raithel (2018) argue that, without scrutiny from investors or capital lenders, self-interested managers who wish to protect their interest and career outlook try to impede the disclosure of unethical corporate behavior by concealing negative information concerning the company. Firms with high levels of long-term debt obligations are monitored more closely by capital lenders (see, e.g., Harris and Raviv, 1990). This fosters the disclosure of unethical corporate behavior.

Risk and return. Prior literature (see, e.g., Gillan and Starks, 2000; Strickland et al., 1996; Karpoff et al., 1996; Ertimur et al., 2011) argues that poor financial performance directly affects the propensity for and the success rate of shareholder activism (i.e., shareholder proposals at the annual general meeting). Hence, we directly address a firm's risk and return dimension, because companies need to adapt their behavior to this kind of shareholder pressure (Oliver, 1991). Otherwise they risk losing their legitimacy and threaten their economic survival.

Consequently, in the aftermath of a poor prior financial performance, myopic profit-seeking shareholders may pressure firms to abandon costly ethical standards and force them to engage in profit-enhancing, less ethical practices, such as cutting payments to workers or reducing workplace safety standards (Campbell, 2007). In contrast, value-oriented shareholders who advocate the long-term value creation of sustainability will use environmental or social shareholder proposals to pressure firms to adhere to ethical norms and values (Shackleton et al., 2022).

Furthermore, corporate controversies pose a major threat to organizational legitimacy, especially for riskier companies with already high variability in earnings or share prices. Companies accused of unethical behavior are often sued and face uncertain future legal claims (Hong and Kacperczyk, 2009; Waddock and Graves, 1997), which in turn further increases their risk. So on the one hand, these companies should act more ethically, whereas on the other, riskier companies are scrutinized more closely by shareholders, which facilitates the disclosure of unethical behavior.

Kacperczyk, 2009; Merton, 1987), whereas others apply a risk perspective as corporate controversies increase the information asymmetry between investors or lenders and managers (see, e.g., Bowen et al., 2008; Dhaliwal et al., 2011; El Ghoul et al., 2011; Robinson et al., 2011; Goss and Roberts, 2011; Sharfman and Fernando, 2008; Ye and Zhang, 2011).

4.3 Data and methodology

4.3.1 Sample and data collection

In Table 4.1, we present the variables employed in our empirical investigation. Here, we also summarize the expected relations of unethical behavior and its disclosure based on the theoretical considerations from the previous section and provide an explicit overall expectation for each variable. We derive the variables from various data sources, mainly Refinitiv Eikon, Datastream, WorldScope, and World Bank. All company-related variables that are dependent on currency are converted into US dollars. Delisted or insolvent companies are considered until the last available rating or financial information. Thus, our results are not influenced by a potential survivorship bias. To ensure comparability, the firm-specific nation-level determinants each refer to the country in which the firm is headquartered.

Dependent variables: measurement of corporate controversies. To investigate and measure firm controversies we use two metrics that capture negative news in global media sources: the absolute *number of controversies*, as a raw count variable, and the Refinitiv *ESG controversies* score, which represents a rating methodology that incorporates these number of controversies and that is benchmarked in relation to the respective industry group. Thus, it offers an easy and comfortable opportunity for investors and researchers to compare and evaluate the level of unethical behavior of companies.

The absolute number of controversies variable considers 23 ESG controversies topics such as privacy or business-ethics related controversies, which indicate whether controversies affect a firm concerning the respective topic and, if so, in how many cases. In particular, this variable is calculated for every firm by using the controversies topics count on an annual basis and therefore measures the exact number of controversies for each company per year. If a scandal affects several controversies topics, then it is accounted for in the controversies count of each of these topics. Thus, the more serious a scandal, the higher is the aggregated number of counts from the controversies topics, which captures the severity and magnitude of a scandal. As a concrete example, we consider VW's Dieselgate: the absolute number of controversies of Volkswagen reveals a peak of 109 controversies in 2016, when this scandal was mainly reviewed.

In this work, the main variable of interest is the ESG controversies score. This score is based on the absolute number of controversies and calculated as an inverse percentile ranking within the respective industry (Refinitiv, 2021). Therefore, it inherits the property of measuring the severity and magnitude of a scandal from the controversies count variable. If a scandal occurs, this decreases the score of the company involved. In contrast to minor scandals, that affect only one fiscal year, high-profile scandals with ongoing legislation

Table 4.1: Definitions, measurements, data sources of country, and company-level determinants.

Category	Variable	Measurement	Source	UB	SD	OE
		Country-level determinants				
Political & legal system	Legislative and corruption	Evaluates regulatory quality, government effectiveness, and absence of corruption	World Bank	−	+	±
	Political participation	Measurement of voice and accountability	World Bank	−	+	±
	Political stability	Likelihood of destabilization or overthrowal of a government	World Bank	−	+	±
	World Press Freedom index (WPI)	Measurement for the degree of freedom of the press	RSF Reporters without borders	⋆	+	+
Cultural system	Power distance index	Extent to which unequally distributed power is accepted and expected	Hofstede (2001); Hofstede et al. (2010)	−	−	−
	Individualism vs. collectivism	Integration of individuals into social groups	Hofstede (2001); Hofstede et al. (2010)	+	+	+
	Uncertainty avoidance index	Social tolerance for ambiguous situations	Hofstede (2001); Hofstede et al. (2010)	−	⋆	−
	Masculinity vs. femininity	Measure for masculinity with regard to gender-specific role patterns within a cultural community	Hofstede (2001); Hofstede et al. (2010)	+	⋆	+
Labor & education system	Skilled labor	Availability of skilled labor in the country concerned	IMD World Competitiveness Report	+	⋆	+
Financial & economic system	Herfindahl-Hirschman index (HHI)	Measurement of nation-level market competitiveness	World Bank	±	⋆	±
	Number of listed companies (LC)	A yearly number of firms listed on the stock exchange in the respective country	The Global Economy	±	⋆	±
	Human Development index (HDI)	Criterion for evaluating the level of development and prosperity of a country	Human Development Report	−	+	±
	Gross domestic product (GDP)	Annual growth rate of the gross domestic product	Datastream	+	−	±
	KOF Globalisation index (KOFGI)	Indicates a country's degree of globalization with regard to economic, social, and political dimensions	KOF Swiss Economic Institute	−	⋆	−
		Company-level determinants				
CSR reputation	ESG score	Environmental, social, governance performance	Datastream	+	+	+
Firm size & visibility	Size	Logarithm of total assets	Datastream	+	+	+
	Analyst coverage	Total number of analysts providing forecasts regarding earnings per share	I/B/E/S	⋆	+	+
Dependency on capital markets	Cash	The sum of cash and short-term investments divided by total assets	Datastream	+	⋆	+
	Leverage	Long-term debt to total assets ratio	Datastream	−	+	±
	Capex	Capital expenditure divided by total assets times 100	Datastream	+	⋆	+
Risk & return	Return on assets (ROA)	Earnings before interest, taxes, and depreciation over total assets	Datastream	±	⋆	±
	Earning variability	Standard deviation of net income before extra items/preferred dividends of the previous five years over total assets	Datastream	−	+	±
	Price volatility	Average annual stock price movement to a high and low from a mean price for each year	Datastream	−	+	±

This table reports definitions, measurements, data sources of country, and company-level determinants as well as a summary of priori expectations for the relation between the employed variables and unethical behavior (UB), societal disclosure (SD), as well as the resulting overall expectation (OE). While for the occurrence of a concrete scandal both prerequisites (the unethical behavior and the disclosure) need to be fulfilled, the tendency towards controversies can also be increased by a specific factor that only affects one of the two legs as (c.p.) there are always other influences that boost the other leg. Sometimes the effect on one leg may be different from the other, so that no clear overall expectations can be concluded.

disputes and lawsuits may also affect the ensuing years and are then accounted for in the scores of later years. Again, a prominent example is provided by the Volkswagen emissions scandal, which led to a sharp drop in VW's controversies score around 2015, and the score also lingered on a low level for the following years.

To allow an intuitive approach, so that high values represent a high controversy density, we rescale the Refinitiv ESG controversies score according to

$$\text{Rescaled ESG controversies score} = 100 - \text{Refinitiv ESG controversies score}.$$

Whenever the ESG controversies score is mentioned in the following, we refer to the rescaled ESG controversies score.

In this paper, we follow a slightly adapted version of Whitley's (1999) well-established framework and account for several dimensions of political, cultural, labor market-oriented, and economic influences. We complement this framework with the inclusion of several important firm-specific characteristics. In the following we describe all national-level and firm-specific variables in more detail.

Independent variables: nation-level determinants. To address the country-specific political environment, we add several political factors from World Bank to our dataset. These variables measure aspects of a country's political system such as the implementation and enforcement of regulations and laws, the overall level of corruption, and the political stability. Due to the strong level of collinearity, these political factors cannot simultaneously be used in regressions. To avoid statistical problems resulting from such a multicollinearity, we run a principal component analysis (PCA) to determine a set of political factors that can be used in statistical analyses. A detailed description of the political factors as well as the weights of the PCA (see Table 4.9) can be found in the Appendix.

As a result of the PCA, we obtain three variables. The first variable, which we refer to as *Legislative and corruption*, predominantly measures the efforts and effectiveness of a government to formulate, establish and enforce laws, as well as the absence of corruption in a country. The remaining two PCA variables mainly cover aspects of political participation and political stability and are specified as *Political participation* and *Political stability*.

Since corporate controversies are detected by global media sources, we also consider the press freedom variable *World press freedom index (WPI)*. This variable is a measure of the degree of press freedom in a particular country. Here, criteria such as pluralism, media independence, transparency, environment- and self-censorship, and violence against journalists are taken into account.

Besides political and legal system determinants, we incorporate cultural aspects. Even

in more developed countries, the societies differ with regard to ethical standards and definitions of appropriate societal and corporate behavior (Matten and Moon, 2008). In order to quantify such ethical standards and diverse definitions of appropriate societal norms and corporate behavior based on geographical country-specific influences, we use the well-known Hofstede cultural dimensions (Hofstede, 2001; Hofstede et al., 2010). These variables are part of a well-established framework, in which the behavior and norms of societies, organizations or a wider range of stakeholders can be understood within the context of national culture (Williams and Zinkin, 2008; Liang and Renneboog, 2017). These dimensions are designed to explicitly examine cross-cultural differences (Beekun and Westerman, 2012) and are therefore perfectly suited to cross-country analyses. In contrast to all other variables, these cultural dimensions are time-invariant. Hofstede's cultural dimensions display long-term cultural developments and are not subject to short-term changes in, for example, distribution of power in society. Moreover, these dimensions should not be perceived as an absolute value, but rather as a cross-country ranking system depicting cultural differences. Specifically, the following four cultural dimensions are utilized.

The *Power distance index (PDI)* variable indicates the degree to which the power imbalances within organizations are accepted and expected in a society (Hofstede and McCrae, 2004). Low PDI stands for open discussion, criticism, and freedom of expression, whereas in the case of a high PDI, neither criticism nor questioning of authority is desired (Tyler et al., 2000).

The integration of individuals into (social) groups is described by the factor *Individualism vs. collectivism (IDV)*. Here, strongly individualistic nations display a stronger self-orientation of society and a decrease in ethical responsibility, whereas in collectivist societies, the welfare of the group is paramount (Ho et al., 2012; Hofstede and McCrae, 2004). The higher the value, the more individualistic the society.

The *Uncertainty avoidance index (UAI)* factor describes the social tolerance toward ambiguous situations. The higher its score, the more the respective society tries to avoid ambiguity and unknown situations and tends to implement stable social norms.

As a proxy for the influences of gender-specific role patterns, we add Hofstede's *Masculinity vs. femininity (MAS)*. Within masculine cultures, power, assertiveness, and success are considered to be particularly important, whereas feminine cultures focus on values such as social support, caring, and helpfulness (Ho et al., 2012).

As a measurement for the labor & education system, we include the *Skilled labor* variable. This variable evaluates a country's education system by assessing the availability of skilled labor. The higher the value, the better the availability of skilled human capital.

Finally, to address the effects of financial and economic systems, we first include the *Herfindahl-Hirschman index (HHI)* and the *Number of publicly listed companies (LC)*.

These variables measure the competitiveness of a country's economy or financial market.

Furthermore, we incorporate the *Human development index (HDI)* as a criterion to evaluate the level of development and prosperity of a country. This variable comprises aspects such as the life expectancy and the general level of of social welfare (Gomanee et al., 2005). Next, we add he annual growth rate of the *Gross domestic product (GDP)* to capture the level of national economic development.

Additionally, we include the *KOF globalisation index (KOFGI)* as a measure of globalization, which may influence corporate controversies, due to a faster flow of information and global networking.

Independent variables: company-level determinants. To cover effects on the company level, we also examine the following firm-related variables. The *Refinitiv ESG* score evaluates a company's environmental, social, and governance performance with regard to 10 main themes (i.e., resource utilization, innovation, emissions, human rights, workforce, management) based on publicly available company-reported data. These data result in the three so-called pillar scores with different weightings (one for each ESG segment) and are finally summarized in the percentile ranked ESG score, which is benchmarked against industry (Refinitiv, 2021). Moreover, this variable is a suitable measure for a firm's CSR reputation.

To examine the influence of firm size we use the variable *Size*, which represents the logarithm of total assets. As both a proxy for visibility and to detect high-attention firms, the number of analysts covering and rating the respective company (*Analyst coverage*) is considered.

To quantify the extent to which companies depend on capital markets, we include the following three variables. On the one hand, we add the variables *Cash*, which consists of the sum of cash and short-term investments in relation to total assets as well as *Capital expenditures (Capex)*, which therefore consequently indicate slack financial resources and a low level of capital market dependency. On the other hand, we add the variable *Leverage* (long-term debt divided by total assets), which comprises debt obligations as an indicator for a rather high dependency on the capital market.

Moreover, to cover aspects of the risk and return dimension, we use *Return on assets (ROA)*, as a measure for firm performance, as well as the risk measures *Earnings variability* and *Price volatility*.

For computational reasons, some variables need to be rescaled. In accordance with the corresponding literature, we winsorize the variables *Size, Cash, Leverage, Capex, Earnings variability, Price volatility*, and *Return on assets* on a 1% level.

4.3.2 Summary Statistics

Our global dataset covers a total of more than 40,000 annual observations for 5,700 different companies located in 44 countries based on the time period between 2002 and 2017. The largest group of the observed firms (roughly one third) are located in the USA, but a large number of companies also come from the UK, Australia, Japan, Canada, China, Hong Kong, and additional European countries[9]. Table 4.2 reports the distribution of observations per year. As Refinitiv expands its data universe, the number of firms in our sample increases over time. On average, our sample covers 2,500 companies per year.

Table 4.2: Number of firm observations per year.

Year	2002	2003	2004	2005	2006	2007	2008	2009
N	688	720	1,337	1,704	1,735	1,851	2,198	2,561
Year	2010	2011	2012	2013	2014	2015	2016	2017
N	3,004	3,120	3,231	3,209	3,283	3,430	3,934	4,739

This table reports the number of observations per year in our dataset.

Table 4.3 shows the descriptive statistics for all variables in our data universe. Concerning the controversies rating, the average value of the rating universe corresponds to about 50 with a standard deviation of approximately 20. The absolute number of controversies exhibits a mean value of 0.45 with a standard deviation of 1.76. Table 4.4 reports the respective correlations coefficients[10]. In Table 4.5 we present the controversies scores grouped by our independent variables.

[9]For a detailed list of number of companies per country see Table 4.10 in the Appendix.

[10]To test for multicollinearity in our dataset, we estimate the variance inflation factor (VIF) values based on an OLS model, which indicates no linear relations for any of our variables.

Table 4.3: Descriptive statistics.

Category	Variable	Mean	St. Dev.	Min	Median	Max
Dependent variable	Controversies score	50.36	20.09	12.50	42.08	99.87
	Number of controversies	0.45	1.76	0.00	0.00	109.00
Political & legal system	Legislative and corruption	0.66	0.75	−2.82	0.84	2.44
	Political participation	−0.10	0.62	−3.43	−0.03	1.49
	Political stability	0.35	0.58	−2.82	0.46	1.12
	World press freedom index	0.75	0.20	0.02	0.79	1.00
Cultural system	Power distance index	46.92	14.64	11.00	40.00	104.00
	Individualism vs. collectivism	70.10	24.61	13.00	80.00	91.00
	Uncertainty avoidance index	55.33	20.77	8.00	46.00	112.00
	Masculinity and femininity	61.47	17.02	5.00	62.00	95.00
Labor & education system	Skilled labor	6.14	0.92	1.88	6.31	8.38
Financial & economic system	HHI	10.46	12.18	3.30	6.08	70.82
	LC	2.79	1.70	0.01	3.28	5.84
	HDI	88.49	6.36	55.80	90.60	95.30
	GDP	2.13	2.31	−9.13	2.22	25.16
	KOFGI	79.91	6.84	54.51	81.17	91.31
CSR reputation	ESG Score	50.83	17.28	5.06	50.08	97.66
Firm size & visibility	Size	15.36	1.49	10.97	15.34	19.34
	Analyst coverage	12.10	8.21	0.00	11.00	57.00
Dependency on capital markets	Cash	0.14	0.14	0.00	0.09	0.82
	Leverage	0.21	0.17	0.00	0.19	0.89
	Capex	5.81	6.08	0.00	4.06	40.69
Risk & return	Return on assets	0.11	0.11	−0.60	0.11	0.49
	Earnings variability	0.46	0.69	0.01	0.23	6.09
	Price volatility	28.87	9.87	10.67	27.24	64.08

This table presents the mean, standard deviation, median, minimum, and maximum values of all variables of the full dataset ($N = 40,744$). All variables are as described in Table 4.1.

Table 4.4: Further descriptive statistics.

Variable		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Controversies score	1.00																							
2	Legislative and corruption	0.04	1.00																						
3	Political participation	−0.08	−0.18	1.00																					
4	Political stability	0.06	−0.05	−0.12	1.00																				
5	World press freedom index	0.05	0.40	0.16	0.72	1.00																			
6	Power distance index	−0.08	−0.49	0.06	−0.58	−0.75	1.00																		
7	Individualism vs. collectivism	0.13	0.43	−0.30	0.58	0.60	−0.78	1.00																	
8	Uncertainty avoidance index	−0.06	−0.47	0.32	0.20	0.04	0.27	−0.39	1.00																
9	Masculinity and femininity	−0.04	−0.09	0.21	−0.06	−0.05	0.03	−0.06	0.37	1.00															
10	Skilled labor	0.06	0.42	−0.05	0.06	0.25	−0.21	0.26	−0.15	−0.04	1.00														
11	Herfindahl-Hirschman index	−0.07	0.03	0.21	0.10	0.15	−0.10	0.04	−0.01	−0.11	0.05	1.00													
12	LC	0.08	0.23	−0.29	0.09	0.05	−0.24	0.47	−0.19	0.30	0.33	0.06	1.00												
13	HDI	0.06	0.76	0.19	0.21	0.63	−0.62	0.48	−0.09	0.05	0.41	0.05	0.19	1.00											
14	Gross domestic product	−0.02	−0.06	−0.15	−0.37	−0.44	0.29	−0.23	−0.26	−0.15	−0.05	0.02	−0.03	−0.32	1.00										
15	KOF Globalisation index	0.08	0.59	−0.03	0.41	0.63	−0.64	0.61	−0.18	−0.23	0.22	0.07	−0.05	0.68	−0.26	1.00									
16	ESG score	0.30	−0.03	−0.04	0.13	0.07	−0.04	0.05	0.07	−0.06	−0.03	−0.08	−0.07	0.01	−0.07	0.12	1.00								
17	Size	0.34	−0.07	0.00	−0.10	−0.12	0.18	−0.16	0.16	0.04	0.07	−0.15	0.05	−0.05	−0.01	−0.09	0.46	1.00							
18	Analyst coverage	0.26	0.01	−0.13	0.01	−0.06	0.07	−0.01	−0.04	−0.10	0.12	−0.20	0.05	−0.03	0.04	0.04	0.40	0.46	1.00						
19	Cash	−0.03	0.02	0.00	−0.09	−0.07	0.05	−0.05	0.00	0.07	0.02	−0.08	0.03	0.02	0.03	−0.07	−0.11	−0.25	0.05	1.00					
20	Leverage	0.05	0.03	−0.10	0.07	0.03	−0.07	0.15	−0.08	−0.08	0.04	0.01	0.11	0.04	−0.04	0.07	0.05	0.20	−0.04	−0.33	1.00				
21	Capex	−0.04	−0.02	0.01	0.03	0.02	−0.03	0.03	−0.04	−0.07	−0.03	0.17	0.02	−0.05	0.06	−0.02	−0.05	−0.08	0.00	−0.12	0.06	1.00			
22	Return on assets	−0.01	−0.04	−0.08	0.02	−0.04	0.01	0.01	−0.04	−0.04	0.00	−0.06	0.01	−0.09	0.07	−0.02	0.07	0.00	0.18	−0.01	−0.09	0.14	1.00		
23	Earnings variability	−0.02	0.08	−0.01	0.08	0.11	−0.14	0.15	−0.11	−0.08	0.02	0.09	0.00	0.09	−0.01	0.08	−0.12	−0.35	−0.13	0.24	−0.04	0.08	−0.29	1.00	
24	Price volatility	−0.07	−0.02	0.00	−0.03	−0.01	0.00	−0.02	−0.08	−0.05	0.02	0.06	0.01	−0.04	0.02	−0.11	−0.24	−0.34	−0.13	0.26	−0.09	0.12	−0.25	0.47	1.00

This table shows the absolute values of correlations.

Table 4.5: Controversies rating grouped by independent variables.

Variable	≤ Median	> Median	Difference
Legislative and corruption	50.27	50.44	−0.17
Political participation	52.45	48.15	4.31***
Political stability	49.86	50.84	−0.98***
World press freedom index	50.09	50.62	−0.53***
Power distance index	52.13	47.26	4.87***
Individualism vs. collectivism	48.10	52.87	−4.77***
Uncertainty avoidance index	52.47	48.06	4.41***
Masculinity and femininity	51.06	48.79	2.27***
Skilled labor	49.07	51.60	−2.54***
Herfindahl-Hirschman index	52.81	47.84	4.97***
LC	48.85	51.95	−3.10***
HDI	48.62	52.08	−3.46***
Gross domestic product	50.22	50.51	−0.30
KOF Globalisation index	50.05	50.65	−0.60***
ESG score	45.57	55.15	−9.58***
Size	45.14	55.60	−10.46***
Analyst coverage	46.57	54.58	−8.00***
Cash	50.56	50.15	0.41**
Leverage	49.44	51.28	−1.84***
Capex	50.35	50.37	−0.02
Return on assets	50.46	50.25	0.21
Earnings variability	50.30	50.41	−0.11
Price volatility	51.51	49.20	2.31

This table reports the mean of Controversies score grouped by each variable. All variables are as described in Table 4.1. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

4.3.3 Methodology

By construction, our data is subject to different structures related to the country and company characteristics (see, i.e., political factors, cultural dimensions, and company-related data) and regarding the data frequency. More precisely, the company- and country-related data, as well as political variables, are calculated on an annual basis, whereas Hofstede's cultural dimensions are time-invariant variables.

Furthermore, to investigate the influences on the occurrence of corporate controversies as accurately as possible, we focus on considering variation over time and between the firms. In order to capture within and between effects in one model, we use a so-called hybrid regression model (see Allison, 2009; Schunck, 2013). This hybrid model, which is also

called within-between regression model, is basically defined by

$$y_{it} = \beta_0 + \beta_1(x_{it} - \bar{x}_i) + \beta_2 c_i + \beta_3 \bar{x}_i + \mu_i + \epsilon_{it} \tag{4.1}$$

where y_{it} denotes the dependent variable for an individual i at time t, x_{it} represents a variable that varies over time and individuals, whereas c_i is a variable that varies only over individuals. Let $\bar{x}_i$ denote the mean of the x_{it} for a fixed i over t. Moreover, μ_i is an error term and random intercept, while ϵ_{it} is considered as a noise variable. Furthermore, using (4.1) we are able to estimate the within effect (β_1) as well as the between effect (β_3) in single models, while keeping time-invariant effects (β_2). Thus, this model allows us to separately measure and interpret variation over time (within) as well as between individuals and is therefore particularly suitable to measure influences on the Controversies score.

As further analysis we also use the number of controversies as dependent variable. Since most of the observed companies are facing none or only a few controversies, our dataset clearly exhibits characteristics of left-censored data. In this context we use a Tobit regression model with clustered standard errors on an industry level (Tobin, 1958). This model estimates the time-dependent relationships between variables, but does not differ regarding within and between effects.

4.4 Results

4.4.1 Regression results for the contemporary regression

In this section, we analyze the contemporary[11] controversies score as a dependent variable. Since the hybrid regression model considers both within and between effects, the results comprise two parts. Here, the within results examine the effect of changes over time, whereas the between results compare the cross-section of firms. Table 4.6 exhibits the results of the hybrid regression.

Country-level determinants. The variable legislative and corruption reveals a negative and significant coefficient at the 1% level in both within and between results. Considering the within effect, this indicates that enforcement of the legislative and a low level of corruption over time lead to a decrease in controversies. Moreover, it appears that the social expectations regarding responsible behavior, set by laws (Aguilera et al., 2007), have a positive impact on the occurrence of corporate controversies, i.e., lead to fewer corporate controversies.

The results of the between effect indicate that companies in countries with strong legislative and corruption factors tend to have significantly lower controversies scores and therefore

[11] Contemporary regression means that the indepenent variables are not lagged.

Table 4.6: Hybrid regression based on contemporary variables.

		Within effects		Between effects	
Category	Variable	Coefficient	S.E.	Coefficient	S.E.
Political & legal system	Legislative and corruption	-2.97^{***}	1.03	-1.56^{***}	0.64
	Political participation	-0.46	0.70	-0.94^{**}	0.49
	Political stability	-3.70^{***}	1.20	0.06	0.66
	World press freedom index	-0.88	1.72	0.90	2.50
Cultural system	Power distance index			-0.04^{**}	0.02
	Individualism vs. collectivism			0.11^{***}	0.01
	Uncertainty avoidance index			-0.06^{***}	0.01
	Masculinity and femininity			-0.01	0.01
Labor & education system	Skilled labor	0.15	0.18	-0.30	0.27
Financial & economic system	HHI	-0.23^{***}	0.07	0.00	0.01
	LC	-0.79^{***}	0.30	-0.06	0.14
	HDI	1.43^{***}	0.12	0.16^{**}	0.07
	GDP	0.09^{*}	0.05	0.34^{**}	0.16
	KOFGI	-0.08	0.12	-0.07^{*}	0.04
CSR reputation	ESG score	0.10^{***}	0.01	0.16^{***}	0.01
Firm size & visibility	Size	1.82^{***}	0.28	4.29^{***}	0.14
	Analyst coverage	0.08^{***}	0.02	0.14^{***}	0.03
Dependency on capital markets	Cash	0.14	1.41	5.26^{***}	1.27
	Leverage	-1.57	1.14	-4.12^{***}	1.00
	Capex	-0.07^{***}	0.02	-0.03	0.03
Risk & return	Return on assets	-7.56^{***}	1.29	2.78	1.74
	Earning variability	0.67^{***}	0.23	1.61^{***}	0.29
	Price volatility	0.06^{**}	0.02	0.09^{***}	0.02
	Pseudo R^2 (total)	0.31			

This table shows the results derived from the within-between regression based on the full sample. The regressions are calculated based on contemporary variables. Coefficients of within-regression (β_1) and between-regression (β_2 and β_3) results, standard errors, and pseudo R^2 are reported upon. All variables are as described in Table 4.1. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

fewer controversies than competitors in less stable countries. One explanation for this could be that in a weakened political environment, unfair competition is more likely to be necessary for the maintenance of economic viability and firms may consider unethical practices with a view to accomplishing competitive advantages or at least to avoiding competitive disadvantages.

The coefficients of the political participation variable reveal a negative and significant value in the between results as well as a negative but insignificant within-coefficient. This may be due to the fact that in countries with a high level of political participation, the disclosure of corporate controversies tends to occur more frequently and therefore companies appear to maintain their ethical reputation more carefully.

Moreover, we observe a negative and strongly significant coefficient of the within effect of the political stability factor as well as a slightly positive but insignificant coefficient of the between effect. A high level of political stability enables members of society to express their concerns and to mobilize others in the direction of activism, which is naturally associated

with particular controversy topics (Cai et al., 2016). Hence, companies try to handle their reputation more responsibly and tend to be less likely to become involved in scandals.

All in all when considering political dimensions, we find evidence that in countries with an efficient enactment and enforcement of laws and low levels of corruption, high levels of political participation and stability, companies are less likely to become involved in corporate controversies. To put this finding in a nutshell: The stronger a political and legal system, the less companies within it are involved in corporate controversies. Besides, since neither the within- nor the between-results of the world press freedom index variable indicates a significant coefficient, we find no indication for a relation to corporate controversies.

Whilst considering the power distance index as a measure for the unequal distribution of power within societal structures, we observe supporting evidence in favor of the a priori overall expectation. The coefficient of the power distance index variable is negative and significant at the 5% level, which illustrates that there is a negative relationship between a company's controversies score and the level of power distance in the country in which the firm is headquartered. One possible explanation could be that even if decision-makers were able to misuse their power for the pursuit of personal benefits, it is more likely that power is used for far-sighted and well-considered decisions. In addition, they may also fear the possible loss of power associated with uncovered controversies. Alternatively, in line with Hofstede et al. (2010), and Williams and Zinkin (2008), another explanation may be that in high power distance societies, scandals or patterns of unethical behavior of superiors are more likely to be covered up. Their decisions are accepted by individuals and stakeholders without question or the attempt to make any moral judgments. Consequently, this results in a decrease in corporate controversies.

The coefficient of the individualism vs. collectivism variable, as a measure for the level of individualism of a society, is positive and significant at the 1% level. This result illustrates the observation that in societies with higher levels of individualism companies are more likely to be involved in a corporate controversy, which could be explained by the fact that in individualistic societies, people advocate on behalf of their own rights and interests, providing the potential for various controversial topics. As published by Ringov and Zollo (2007), companies which operate in societies with a high level of individualism are pressured into behaving ethically responsibly to a lower degree. The evidence suggests that this results in a higher controversies score and therefore firms tend to be more likely involved in corporate scandals. Altogether, the empirical analysis supports the expected positive relation.

As another cultural dimension, we examine the tolerance toward ambiguity, measured by the UAI variable and observe statistically significant evidence in favor of our negative a priori expectation. This illustrates that companies in countries with a higher level of uncertainty avoidance tend to achieve lower controversies scorings. The reason behind

this finding may be that the UAI variable also indicates how cultures deal with deviations from (entrenched) moral values regarding ethical behavior. Hence, this result underlines the expectation that reactions to deviations from this behavior are far more severe than in cultures with lower UAI, and companies therefore feel coerced into taking care of their social policies.

The coefficient of the masculinity vs. femininity variable indicates a negative but insignificant value which so far does not suggest an impact on corporate controversies.

The coefficients of the availability of skilled labor variable are insignificant and therefore we find no supporting evidence regarding the positive a priori expectation. An explanation for this observation could be the following. Since certain types of scandals are more typical for particular industries and are therefore less strongly perceived, it would also be necessary to distinguish between various types of scandals in order to measure contemporary impacts. Furthermore, from an employee's point of view, the occurrence of media scandals is clearly only one aspect of choice of profession besides other factors such as the availability of equivalent alternative positions, advancement opportunities, or financial aspects, which also play a role for potential applicants.

Furthermore, regarding determinants of the financial system, the Herfindahl-Hirschman index, as a measure of competitiveness, exhibits a negative within-coefficient and an insignificant coefficient regarding the between effect. One explanation for this observation may be that in a business environment with a low level of competitiveness, companies may have little incentives to act ethically. A similar effect can be seen when examining the coefficients of the number of listed companies, which also reveal a negative and significant within effect and negative but insignificant between effect.

The coefficient of the HDI variable is positive and significant in both within and between results. The results illustrate the observation that in more developed societies, companies are more likely to be involved in corporate scandals. One reason for these results could be that controversies are more likely to be detected or systematically uncovered, as they are less greatly accepted in countries with higher levels of prosperity. Note that since all country-specific data refer to the country in which the company headquarter is located, the state of economical and societal development merely applies to the people in the respective country. However, this variable also influences the perception of the business activities of a company in other countries, since potential controversies there—even if they are far more greatly accepted due to local societal demands—are assigned to the company. As an example, one could consider child labor in India while working for a Swedish fashion company, or, to cite a concrete example, the Nestlé child labor controversy in 2005.

The gross domestic product variable displays a significant positive coefficient in both within and between results. This results show that in countries with higher levels of GDP

growth rate, companies are more likely to be involved in corporate scandals. One possible explanation may be that a high GDP growth rate typically can be observed in developing countries with rather low levels of political and social norms. As a result, there is also less pressure on companies to adhere to ethical standards which results in a higher number of corporate controversies.

The KOF Globalisation index variable, exhibit support in favor of a rather negative relationship on a company's controversies score: the coefficient of the between effect is negative and significant, while the within-coefficient is negative but insignificant. Countries with high degrees of globalization may tend to adopt high ethical standards and therefore may exert more pressure on companies to comply with these, which results in an decrease of corporate controversies.

Company-level determinants. In contrast to our nation-level variables, these factors naturally vary over time and between different companies operating in the same country.

When considering the ESG score as an indicator for a company's CSR reputation, we observe positive coefficients in both parts of the hybrid model. The results are significant at the 1% level, which illustrates the positive relationship between the CSR reputation and the controversies score. These findings show that firms associated with good CSR reputations are more likely to become involved in corporate scandals than companies with rather low CSR reputations. Dorfleitner et al. (2020), who examine the relationship between corporate social performance and corporate financial performance, ascertain a similar relationship. Moreover, companies with good ESG values are also expected to have a "clean coat" concerning scandals, but generally fail to meet these expectations. However, one possible reason for this finding could be that these companies tend to rely on the insurance-like effect of high CSR values and fear possible consequences of scandals less. As a result, the incentive to behave ethically also appears to decrease. Altogether, our results are in line with the positive a priori expectation.

The size variable has a strongly significant positive sign in both within and between results, which supports the positive overall expectation, illustrating that a larger firm size leads to a significant increase in corporate controversies and thus a decrease in the controversies score.

As a measurement for a high-attention firm and as an indicator for visibility, we investigate the analyst coverage variable. For this, we observe positive coefficients in within and between results. Both results are significant at the 1% level. The reason behind this observation may be ambiguous. The more analysts examine a company, the easier it is to reveal inconsistencies and to make this information accessible to investors. In summary, our results are in accordance with the positive overall expectation.

Even if the cash variable of the within effect has a positive but insignificant coefficient,

we observe a positive and strongly significant coefficient in the between effects. One explanation for this may be the well-known free cash flow hypothesis (Jensen, 1986) and empire-building theory, which state that companies with a high level of cash and the resulting decrease in dependency on the capital market, tend to be less concerned about their reputation and potential consequences of corporate controversies.

The leverage variable has a negative but insignificant sign regarding the within effects and a negative and strongly significant sign regarding the between effects. Thus, an increase in leverage appears to decrease corporate controversies.

Furthermore, since significant results regarding the cash and leverage variables are only detected from the between results, this illustrates that these effects only occur between companies and changes over time do not show any significant implications. One possible explanation is that the capital structure of a company, which influences the level of dependency on the capital market, changes rather slowly and these changes are generally rather small.

The coefficients of the capex are negative and significant regarding the within effect, and negative but insignificant regarding the between effect. Apparently, higher capital expenditures and the therefore higher capital requirements over time tend to be connected with more responsible corporate behavior. Companies with high capex values tend to be relatively future oriented and consequently place value on their public perception, stakeholder relations and ethical reputation. Contrarily to our a prior expectation, low capex may not solely indicate a low dependency on capital markets but also reflect the lack of sufficient growth and investment opportunities (Ferrell et al., 2016).

Summarizing, there is a negative and significant relationship between return on assets and the controversies score. This illustrates that better-performing firms (measured by ROA) tend to achieve better controversies scores. The coefficients of the two risk variables price volatility and earnings variability indicate a positive and significant association. Hence, riskier firms also tend to reveal more controversies. Of course, there may also be unobserved variables which trigger risk and unethical behavior at the same time. Thus, special care should be applied in interpreting the risk and return finding.

4.4.2 Regression results for future controversies score

While the findings above are very indicative, from an investor's point of view it is much more interesting to ascertain whether and to what extent there is a possibility to predict future corporate controversies. Table 4.7 exhibits the results of the hybrid regression, which analyzes the influence of our explanatory variables in t on the controversies scores in the following year $t + 1$.

Interestingly, the results of the between part of the regression are materially the same as those of the contemporary results. Concerning the within-part of our investigation, the level of significance partly decreases, but the results largely remain unchanged.

Table 4.7: Hybrid regression - future controversies score.

		Within effects		Between effects	
Category	Variable	Coefficient	S.E.	Coefficient	S.E.
Political & legal system	Legislative and corruption	0.27	1.09	−1.82***	0.67
	Political participation	0.58	0.74	−1.11**	0.52
	Political stability	−3.30***	1.26	−0.15	0.69
	World press freedom index	1.87	1.80	0.45	2.65
Cultural system	Power distance index			−0.05**	0.02
	Individualism vs. collectivism			0.12***	0.02
	Uncertainty avoidance index			−0.06***	0.01
	Masculinity and femininity			−0.02	0.01
Labor & education system	Skilled labor	−0.33*	0.19	−0.39	0.29
Financial & economic system	HHI	−0.24***	0.07	0.00	0.01
	LC	−1.04***	0.32	−0.03	0.15
	HDI	1.39***	0.12	0.15**	0.07
	GDP	−0.15***	0.05	0.25	0.17
	KOFGI	−0.14	0.12	−0.08*	0.04
CSR reputation	ESG score	0.08***	0.01	0.16***	0.01
Firm size & visibility	Size	2.38***	0.30	4.37***	0.15
	Analyst coverage	0.02	0.03	0.17***	0.03
Dependency on capital markets	Cash	−0.86	1.49	4.95***	1.35
	Leverage	−0.68	1.21	−4.71***	1.08
	Capex	−0.01	0.03	−0.05	0.03
Risk & return	Return on assets	−3.56***	1.38	2.47	1.88
	Earnings variability	0.55**	0.25	1.64***	0.32
	Price volatility	0.02	0.03	0.11***	0.02
	Pseudo R^2 (total)	0.31			

This table shows the results derived from the within-between regression based on the full sample. Coefficients of within-regression (β_1) and between-regression (β_2 and β_3) results, standard errors, and pseudo R^2 are reported upon. All variables are as described in Table 4.1. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

4.4.3 Regression results for the absolute number of controversies

The calculations of the Tobit regressions based on the contemporary number of controversies as dependent variable are presented in Table 4.11. As the reference industry[12] we choose manufacturing, which is the most frequently represented industry in our dataset.

In comparison with our reference category, only retail trade and agriculture, forestry, and fishing exhibit significantly positive coefficients, whereas the coefficients of the remaining industry dummies indicate a significant negative sign. This points to the fact that in the

[12]The industrial affiliation is based on the SIC manual of the United States Department of Labor respectively.

manufacturing, agriculture, forestry, and fishing, and retail trade industry scandals tend to occur most frequently while simultaneously accounting for other influential variables.

Moreover, the data suggest that almost all previous observations can also be confirmed in this approach. With regard to the nation-level determinants legislative and corruption, political participation, power distance, individualism vs. collectivism, uncertainty avoidance, the results reveal a consistency with our former findings and therefore support these expectations. Even if we observe a positive and significant coefficient of the skilled labor variable, we cannot draw a clear conclusion based on the previous results and therefore still reject to commit to a positive or negative overall relation.

For company-related characteristics, we also find, apart from isolated deviations, a strong link to the previous results. The variables ESG score, firm size, analyst coverage, cash, leverage, earnings variability, and price volatility show highly significant values in the expected direction and therefore, once again, confirm our results.

To verify the differences between the remaining sectors, we again calculate the Tobit regression and vary the reference industry, with all other coefficients remaining unchanged. On the one hand, one can observe that the financial sector as well as the construction sector are generally inclined towards fewer controversies. On the other hand, both the agriculture, forestry, and fishing, and retail trade industries appear to have a strong level of involvement in controversies. This even has a practical impact and can be very useful for an investor who wishes to invest in industries with few controversies or who wishes to avoid potential risks from upcoming controversies.

4.5 Robustness checks

To examine the robustness of our results, we conduct various further computations. Since several companies operate their businesses in multiple countries, they are also far more greatly affected by various nation-level influences. To take this into account, we add a multinationality index in order to measure to which extent a company operates multinationally. We define companies with more than 10% of international assets, i.e., foreign assets divided by total assets, as being "multinational". In line with Ioannou and Serafeim (2012), we split our sample to distinguish between domestic (this includes approximately 19,400 observations) and multinational firms. For both samples, we run the hybrid and Tobit regressions again. The results of the domestic and multinational regressions are displayed in Table 4.8 for the hybrid regression.

When considering the statistics of the contemporary results, we observe that most major results are confirmed in both samples. This affects all company-level determinants regarding CSR reputation, size, cash, leverage, capex, return on assets, earnings variability, and

Table 4.8: Hybrid regression - domestic & multinational companies (contemporary variables).

	Domestic companies				Multinational companies			
	Within effects		Between effects		Within effects		Between effects	
Variable	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Legislative and corruption	−5.78***	1.51	−1.68*	0.89	0.93	1.48	−1.12	1.02
Political participation	−2.32**	1.01	−1.10	0.73	2.15**	1.01	−0.08	0.75
Political stability	−6.15***	1.76	−0.73	0.81	−0.76	1.69	2.19	1.35
World press freedom index	−2.45	2.36	0.60	3.60	2.90	2.62	−2.51	3.98
Power distance index			−0.07**	0.03			−0.03	0.03
Individualism vs. collectivism			0.10***	0.02			0.12***	0.02
Uncertainty avoidance index			−0.05**	0.02			−0.07***	0.02
Masculinity and femininity			−0.03	0.02			−0.00	0.02
Skilled labor	0.23	0.27	0.28	0.39	−0.01	0.25	−0.44	0.43
HHI	−0.34***	0.11	−0.02	0.02	−0.23**	0.09	0.02	0.02
LC	−1.13***	0.42	0.14	0.19	−0.37	0.45	−0.26	0.22
HDI	1.76***	0.18	0.06	0.10	1.04***	0.17	0.25**	0.10
GDP	0.12*	0.07	0.15	0.22	0.11*	0.07	0.51**	0.26
KOFGI	−0.31*	0.18	−0.02	0.06	0.14	0.17	−0.16***	0.06
ESG score	0.07***	0.02	0.12***	0.01	0.12***	0.02	0.19***	0.02
Size	1.85***	0.40	3.52***	0.20	2.05***	0.40	5.03***	0.23
Analyst coverage	−0.01	0.04	0.13***	0.03	0.15***	0.03	0.14***	0.04
Cash	−0.77	1.89	3.15**	1.51	1.20	2.15	8.63***	2.42
Leverage	−0.23	1.59	−3.90***	1.21	−2.09	1.68	−4.28**	1.80
Capex	−0.06*	0.03	−0.05	0.04	−0.09**	0.04	−0.04	0.05
Return on assets	−6.64***	1.76	0.62	2.07	−7.57***	1.95	5.25	3.26
Earnings variability	0.66**	0.31	0.96***	0.37	0.91**	0.36	2.08***	0.49
Price volatility	0.06*	0.03	0.10***	0.03	0.03	0.04	0.08***	0.03
Pseudo R^2 (total)	0.28				0.28			

This table shows the results derived from the within-between regression based on the subsample of domestic or multinational companies. Coefficients of within-regression (β_1) and between-regression (β_2 and β_3) results, standard errors, and pseudo R^2 are reported upon. All variables are as described in Table 4.1. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

price volatility.

However, as expectable, the main differences now lie in the influences of country-level variables. As anticipated for multinational companies, the significance of political factors decreases or changes, which we attribute to the influence of the partially differing political structures between the various countries. Nevertheless, the results of financial and economic system determinants remain by and large in line with previous findings.

Furthermore, an application of an even more restrictive variant of setting the threshold of foreign assets to distinguish between "domestic" and "multinational" to 0%, the results remain materially the same, thus eliminating the need for us to publish them in this paper.

As another robustness test we calculate both, the Tobit regression for the dependent variable *logarithmic number of controversies plus one* as well as for the *number of controver-*

sies one year later by lagging the independent variables, thereby performing a forecasting analysis. The results remain, apart from some isolated deviations from the main model, qualitatively the same and therefore again confirm prior findings. Hence, we do not report them in this work.

Besides that, we also run an OLS regression with clustered standard errors on firm-level as well as fixed effects regressions, which can partly eliminate time-invariant aspects of firm-level endogeneity. Again, the results remain by and large in line with previous findings.

Last but not least, we calculate the domestic and multinational regressions on the future controversies score and the future number of controversies for both hybrid and Tobit regressions. Except for a few deviations, the results remain in line with previous findings and are therefore not reported here either.

4.6 Conclusion

In this paper, which is the first to empirically investigate a broad spectrum of drivers of corporate controversies, we examine an extensive international dataset including over 5,700 companies in 44 countries and the associated measurements for social irresponsibility, as measured by a controversies score and absolute number of controversies, as well as further country-level and company-level variables in the investigation period from 2002 to 2017. We provide empirical insights supporting the significant role of diverse variables in explaining corporate controversies.

Based on extensive research of prior literature, we argue that country-level variables, regarding aspects of the political and legal system, culture, labor and education, and financial and economic system, as well as company-level determinants which capture CSR reputation, firm size and visibility, dependency on capital markets, and risk and return dimension impact the occurrence of corporate controversies.

To examine empirical evidence of our a priori expectations, the calculations are based on the within-between hybrid regression model as well as a Tobit regression model, both of which show that we can identify various factors that reduce the tendency of companies becoming involved in scandals.

In general, the occurrence of corporate controversies comprises a two-stage process: first, the unethical corporate behavior of a firm, and second, the process of societal disclosure, including the perception, disapproval and publication of this unethical behavior. We find evidence that companies will adopt ethical standards and therefore are less likely to become involved in scandals, if they operate in an environment in which they feel high levels of institutional pressure (institutional theory) or in which corporate controversies pose a high threat to organizational legitimacy (legitimacy theory). Moreover, societies with

high moral standards that closely monitor corporate behavior tend to uncover corporate controversies more frequently.

Regarding country-level determinants, on the one hand, particularly in countries with efficient enactment of laws and law enforcement as well as low levels of corruption, high degrees of political participation, political stability, societies with higher power distance, in uncertainty avoiding societies, companies are less likely to be involved in a corporate controversy. On the other hand, in countries with an individualistic culture, in countries with high levels of development and prosperity, as well as high GDP growth, companies are more likely to become involved in a corporate controversy.

Considering company-specific factors, firms associated with good CSR reputation, larger firms and high-attention firms, as well as riskier firms are more likely to become susceptible to corporate controversies, whereas companies with a high level of dependency on capital markets are less likely to become involved in corporate controversies.

Furthermore, our study has practical implications. First of all, many influential factors related to the controversies score are intuitively comprehensible and can easily be adopted by strategists. Thus, these findings can immediately be employed for the implementation of various investment strategies for both private and professional investors. In addition, our results are of benefit to ethically motivated managers, who may find implications for their evaluation and decision-making processes.

One potential limitation of our investigation lies in the fact that both the controversies score and the absolute number of controversies, are only calculated once per annum. Future research could focus on a higher evaluation frequency, which allows a closer investigation of short- and long-term effects, as well as different stock market reactions based on various controversy topics such as environmental controversies or social controversies. What is more, the influence of industry sectors on the occurrence of corporate controversies beyond our approach may also merit a closer examination. Lastly, it must be noted that our research design provides correlational evidence for our theoretical model, which is surely a proper first step. However, in order to provide more causal evidence for the results, deeper insights into the motivations of the involved actors appear to be necessary, which could be achieved by conducting experimental or interview-based research. For some firm-level variables, such as the risk measures earnings variability and price volatility, even a reverse causal relationship would be conceivable. In this paper, however, we do not intend to discuss the causal relationship for each dimension and, therefore, we leave this task to future research.

All in all, our findings underline the key roles which firm- and industry-level factors play, as well as cultural and political influences, in the quest to examine what drives corporate controversies. Moreover, the number and frequency of corporate controversies as a mea-

sure of social irresponsibility, as well as a new dimension of ESG, still appear to harbor promising potential for further research.

4.7 Appendix

Political factors To reflect country-specific political settings and effects and to investigate their potential impact on further investigations as accurately as possible, we add several political dimensions to our dataset through the inclusion of various political factors from World Bank.

To estimate the ability of a government to formulate and establish sound policies and regulations regarding the promotion and permission of private sector development (Kaufmann et al., 2010), we use the *Regulatory quality* variable. Furthermore, *Control of corruption* indicates how far public power is exercised for all forms of corruption and assertion of interests by the elite.

Political stability and absence of violence/terrorism—in the following short *Political stability*—measures the likelihood of destabilization or toppling of a government, also including terrorist activities or politically-motivated violence (Kaufmann et al., 2010). Moreover, the variable *Voice and accountability* captures the extent to which citizens are able to participate in the selection of their government and the degree of freedom with regard to expression, association, and the media. To evaluate the degree to which agents comply with and rely on the rules of society, and in particular the quality of contract enforcement, the police, property rights, and the courts, as well as the likelihood of crime and violence (Kaufmann et al., 2010), we add the *Rule of law* variable. Last but not least the variable *Government effectiveness* depicts the quality of civil service as well as policy formulation and the degree of its independence from politically motivated pressures, the quality of policy implementation and formulation, and the level of credibility of the government's commitment to these policies (Kaufmann et al., 2010).

Table 4.9: Weights of the PCA.

Variables	Legislative and corruption	Political participation	Political stability
Voice and accountability	−0.57	1.39	−0.11
Political stability	−0.71	−0.12	1.58
Government effectiveness	0.53	−0.18	−0.27
Regulatory quality	0.51	−0.10	−0.33
Rule of law	0.35	−0.09	−0.07
Control of corruption	0.49	−0.21	−0.17

This table shows the respective weights of the PCA to calculate the three political dimensions Legislative and corruption, Political participation, and Political stability.

Table 4.10: Number of companies per country.

Country	N	Country	N
United States	13323	Finland	225
Japan	4854	Indonesia	224
United Kingdom	3305	Mexico	224
Canada	2762	Chile	198
Australia	2657	Greece	189
Hong Kong	1437	New Zealand	187
France	1159	Thailand	186
Germany	974	Denmark	183
South Korea	803	Austria	160
South Africa	763	Philippines	158
China	757	Turkey	156
Switzerland	717	Poland	146
India	623	Ireland	144
Brazil	612	Portugal	111
Singapore	531	Israel	102
Sweden	481	Colombia	68
Spain	438	Peru	66
Netherlands	410	Argentina	48
Italy	368	United Arab Emirates	33
Belgium	310	Hungary	27
Malaysia	299	Czech Republic	24
Norway	287	Saudi Arabia	15

This table shows the number of companies per country in the data sample.

Table 4.11: Tobit regression - contemporary variables.

Category	Variable	Coefficient	S.E.
Industry	Agriculture, Forestry, and Fishing	1.98***	0.20
	Construction	−1.66***	0.17
	Finance, Insurance, and Real Estate	−2.79***	0.21
	Mining	−0.70***	0.15
	Retail Trade	0.63***	0.08
	Services	−0.23**	0.10
	Transport, Communications, Electric, Gas, and Sanitary service	−0.34***	0.09
	Wholesale Trade	−0.97***	0.08
Political & legal system	Legislative and corruption	−0.86***	0.24
	Political participation	−0.29***	0.08
	Political stability	0.08	0.25
	World press freedom index	−0.16	0.80
Cultural system	Power distance index	−0.03***	0.01
	Individualism vs. collectivism	0.03***	0.01
	Uncertainty avoidance index	−0.02***	0.00
	Masculinity and femininity	0.00	0.00
Labor & education system	Skilled labor	0.25**	0.10
Financial & economic system	HHI	0.00	0.01
	LC	−0.01	0.08
	HDI	0.01	0.02
	GDP	0.00	0.03
	KOFGI	0.02	0.02
CSR reputation	ESG score	0.04***	0.00
Firm size & visibility	Size	2.09***	0.18
	Analyst coverage	0.04***	0.01
Dependency on capital markets	Cash	2.93***	0.46
	Leverage	−1.59***	0.29
	Capex	−0.02	0.02
Risk & return	Return on assets	1.94***	0.59
	Earnings variability	0.46***	0.09
	Price volatility	0.04***	0.01
	Pseudo R^2	0.15	

This table shows the results derived from the Tobit regression based on the full sample. The dependent variable is the absolute number of controversies, and hence, left-censored at 0. All variables are as described in Table 4.1. The McFadden pseudo R^2 is reported upon. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Chapter 5

Board responsibility for irresponsibility: The influence of board structure on corporate scandals

This research project is joint work with Gregor Dorfleitner (University of Regensburg).

Abstract Based on an international dataset that comprises over 6,100 companies located in 44 countries in the years 2002–2018, this paper analyzes the relation between socially irresponsible behavior and board structures besides further firm-related, political, nation-level economical, and cultural variables. We identify board structure variables that tend to increase a firm's irresponsible corporate behavior, namely high CSR efforts and busy board members. There are also variables that tend to reduce this kind of behavior, namely qualified and skilled boards and larger boards. No clear evidence can be determined from a board's gender diversity.

Keywords Corporate controversies, board structure, corporate scandals, within-between model, hybrid regression model

JEL M14, M12, G30

5.1 Introduction

The term corporate social responsibility (CSR) is attracting increasing attention and has been studied from a practical and academic point of view for many years. Following Liang and Renneboog (2017), CSR is understood as business activities that focus on the improvement of social welfare but not necessarily at the cost of profits or shareholder value. Furthermore being socially responsible entails not only the idea of doing "good" but also includes responsibility for avoiding "bad" in terms of illegal, unethical, as well as social irresponsible behavior (see Lin-Hi and Müller, 2013). Prior work demonstrates that social irresponsible behavior results in direct negative consequences for both, i.e., companies and stakeholders, such as losses in market value (see Karpoff et al., 2005) as well as damage to reputation (see Grappi et al., 2013).

In this regard, Kotchen and Moon (2012) demonstrate that an increase in irresponsible or unethical behavior of companies also tends to lead to an increase in CSR activities, indicating that companies may try to offset corporate irresponsible behavior with social responsible behavior. Moreover, other authors investigate further potential drivers and motivations of companies to engage in CSR (see Ioannou and Serafeim, 2012; Liang and Renneboog, 2017; Reverte, 2009).

In recent years, academic literature has begun to focus on the occurrence of CSR-related controversies[1]. Since some scandals are not only directly linked to economic damage to the respective company but also result in far-reaching and grave environmental damage (see Deepwater Horizon explosion), the topic of corporate ethical behavior concerns economists and is therefore worthy of extensive study.

Following Dorfleitner et al. (2021b), the occurrence of corporate scandals requires two prerequisites: First, the irresponsible or unethical behavior of a company. Second, the process of societal disclosure including the perception, and disapproval as well as the publication of this behavior. Furthermore, the authors identify various political, cultural, as well as country- and firm-related variables that influence companies to become involved in a corporate controversy.

Since many scandals result from unethical or morally questionable decisions by executives, it is necessary to examine the influence of the board, being the entity supervising the executives, on scandals from an academic point of view. More precisely, this work focuses on the influence of board structure variables on the occurrence of corporate controversies.

Moreover, Jain and Zaman (2020) investigate the relation between board structure and corporate social irresponsibility and identify board-level governance conditions to reduce

[1]Note that in everyday language, the notion of controversy still comprises two legitimate opposite perspectives, while the term scandal refers much more to deplorable behavior. Nevertheless, in this article we use both terms interchangeably, which is inspired by the Refinitiv controversies score methodology.

irresponsible behaviors.

This paper differs from their work in several ways.

First, their dataset only comprises US companies and contains merely firm-related variables as control variables. Furthermore, no political, country-related, or cultural variables are considered, which play a big role in explaining corporate social responsibility (see Ioannou and Serafeim, 2012) as well as corporate social irresponsibility (see Dorfleitner et al., 2021b). Second, we capture effects regarding the changes over time as well as cross-section firm effects. Third, Jain and Zaman (2020) measure corporate social irresponsibility based on aggregated costs as well as the total number of unethical incidents, which could be time-consuming to implement for investors and strategists. To avoid such a procedure, we incorporate a scoring methodology to measure firm controversies, which is more easily to adapt.

5.2 Theoretical development

In this section, we hypothesize how board level determinants may reduce or increase the likelihood of a company to be involved in a corporate controversy.

Board Size A company's board holds the authority and responsibility to advise executives and monitor the decision-making process (see Adams and Ferreira, 2007). In accordance with agency theory (see Jensen, 1986), the preferred projects from managers not always reveal maximized shareholder value. Moreover, agency cost theory hypothesizes that managers can make self-serving decisions (i.e., empire building), which decreases shareholder value as a result of decreasing disclosure quality (see Jensen, 1986; Levinson, 2004).

In particular, corporate controversies are often linked to negative implications, like significant price losses as well as ongoing lawsuits and fines and thus harm many of their stakeholder groups (see Fauser and Utz, 2021). Such unpredictable risks are something most investors want to avoid as much as possible.

Previous academic literature finds a negative relation between board size and firm risk (see, e.g., Coles et al., 2008; Wang, 2012). Thus, in terms of corporate controversies, firms less willing to take risks may also exhibit farsighted decision-making process and thus reduce the risk of corporate controversies.

In contrast, the benefits of large boards could be influenced by disadvantages of larger groups regarding poor coordination, less flexibility, as well as bad communication (see De Andres et al., 2005; Lipton and Lorsch, 1992; Jensen, 1993), which could be linked

to an increased occurrence of corporate controversies. Further research also shows that smaller boards tend to be more effective and to reduce free-riding risks (see Yermack, 1996; Ahmed et al., 2006) while others find positive impacts of board size on firm efficiency (Huang et al., 2011).

Regarding CSR activities, prior work reveals evidence for a positive relation of board size and CSR disclosure (see Jizi et al., 2014). Moreover, by running a meta-analysis Zubeltzu-Jaka et al. (2020) find a positive effect of board size on a firm's corporate social performance.

In summary, although the evidence is not conclusive, we expect companies with larger boards to be less frequently involved in scandals.

Hypothesis 1: Companies with larger boards are less likely to be involved in a corporate controversy.

Board expertise In particular financial experts on a board are expected to recognize risks (see Harris and Raviv, 2008) which will not pay off and advise executives to avoid them (see Minton et al., 2014). Thus, board expertise may also affect the occurrence of corporate controversies, such as risks related to cultural or national compliance principles.

Previous literature mainly focuses on the relation of board member education and firm performance (see Darmadi, 2013; Gaur et al., 2015; Bathula, 2008). The results of these studies are quite divergent. While some authors find positive influences (see Darmadi, 2013; Gaur et al., 2015), others find negative tendencies on performance (see Bathula, 2008). Aside from that, Huang et al. (2011) find a negative relation between firm efficiency and the proportion of financial experts on the audit committee.

However, even if board qualification seems to influence firm characteristics in different ways, literature which focuses on the relation between board qualification and corporate social irresponsibility is rare.

As scandals often show unforeseen and long-term effects and therefore even unpredictable risks, we expect a more qualified board with a high level of expertise to avoid the involvement of their company in scandals.

Hypothesis 2: Companies with qualified and skilled boards are less likely to be involved in a corporate controversy.

Board gender diversity Regarding board gender diversity, the majority of the literature focuses on its effect on corporate financial performance (see Rao and Tilt, 2016). However, further strands of literature examine the relation between board gender diversity and corporate social performance and reveal mixed results. Some researchers have

found evidence for a positive relationship (see Post et al., 2011; Boulouta, 2013; Hafsi and Turgut, 2013; Webb, 2004; Mallin and Michelon, 2011), while others found no evidence for a relationship (see Coffey and Wang, 1998; Rodriguez-Dominguez et al., 2009; Manita et al., 2018).

Some authors indicate that women on boards improve decision-making and provide more effort on monitoring (see Adams and Ferreira, 2009). But following Adams and Ferreira (2007), this needs not be an advantage, since tougher monitoring may lead managers to be less willing in sharing information. This lack of communication may result in a decrease in board effectiveness and consequently increase the occurrence of controversies.

However, several authors emphasize that women are in general underrepresented in boardrooms (see Claringbould and Knoppers, 2007; Orbach, 2017; Chapple and Humphrey, 2014). In line with Claringbould and Knoppers (2007), one explanation for women's frequent reluctance to become board members could be based on gender-specific role patterns which are rarely seen as issues for their male counterparts. Another reason may be that females behave more risk-averse than males (see Croson and Gneezy, 2009; Carter et al., 2017), which may also affect the underrepresentation of women on boards.

Corporate social irresponsibility and board diversity is seldom discussed in academic literature. However, early studies find that more diverse boards reduce irresponsible behavior (see Godfrey et al., 2020; Jain and Zaman, 2020). Therefore, we expect firms with high levels of board gender diversity to be less likely involved in corporate controversies.

Hypothesis 3: Firms with high-levels of board gender diversity are less likely to be involved in a corporate controversy.

Busy board members Besides board characteristics like size or expertise, the issue of busy board members, which hold several member affiliations, may also affect the behavior of companies and the involvement in corporate controversies. According to Fich and Shivdasani (2006), there is an inverse relation between busy boards and firm performance, comprising market-to-book ratio as well as operating profitability, when the majority of outside directors hold three or more directorships. However, Field et al. (2013) find evidence that newly public companies benefit from "overboarded" directors since they offer unique advantages in terms of a high level of connection and experience. Ferris et al. (2020) confirm both of these findings and additionally publish that multiple directorships are negatively associated with female directors.

Previous literature also shows that large parts of unethical and eventually illegal activities within a business are attributed to a lack of board oversight (see Murphy and Schlegelmilch, 2013). Additionally, as proposed by Ormiston and Wong (2013), boards should remain vigilant to prevent leaders—if they are high on moral identity symbolization—in particular

to become involved in unethical behaviors. This may become difficult or fail as a result of busy board members.

However, even if a busyness of directors may be a sign of personal expertise (see Fama and Jensen, 1983), a high level of "overboarded" members may lead to a decrease in board efficiency and consequently weaken decision-making processes. In summary, we expect busy boards to encourage the occurrence of a corporate controversy.

Hypothesis 4: Firms with busy board members are more likely to be involved in a corporate controversy.

CSR efforts on board level: CSR committee CSR committees, which are also called sustainability or ethics committees, reveal the willingness of a company to improve its corporate behavior (see Mallin and Michelon, 2011). While most academic literature focuses on the relation between CSR committees and corporate social performance (see Mallin and Michelon, 2011; Eberhardt-Toth, 2017; Baraibar-Diez and D Odriozola, 2019), only few studies investigate the effects between CSR committees and corporate social irresponsbility. Jain and Zaman (2020) as well as Fu et al. (2020) find evidence for a negative relation between the existence of CSR committees and social irresponsible behavior.

Nevertheless, Dorfleitner et al. (2021b) find that companies with high levels of ESG also tend to show high levels of corporate social irresponsible behavior. Furthermore, following Kotchen and Moon (2012), companies may try to offset irresponsible behavior with social responsible activities. Thus, one potential reason for a company to establish a CSR committee may lie in former irresponsible behavior.

Another argument may be that companies with high ESG scores, which quantify the success of CSR efforts, may also be measured by higher standards regarding corporate behavior, making it easier for them to become involved in scandals. Furthermore, a CSR committee may increase the monitoring intensity of managers. As a consequence, in line with Adams and Ferreira (2007), managers may be less inclined to share information, decreasing the ability of a board to monitor effectively.

Taking all aspects into account, arguments for a positive relation outweigh those for a negative relation. Thus, we expect that CSR efforts on the board level are positively related to the occurrence of irresponsible behavior.

Hypothesis 5: Firms with high CSR efforts on the board level are more likely to be involved in a corporate controversy.

5.3 Data and methodology

5.3.1 Sample and Data collection

For our analyses, we use a global dataset by combining various data sources, mainly Refinitiv Eikon, Datastream, WorldScope, and World Bank. It contains information from over 6,100 companies located in 44 countries, based on the time period 2002–2018. The largest percentage of the observed firms are located in the USA (about 36%), but also a large number of companies based in the United Kingdom, Japan, Canada, Australia, Hong Kong, China as well as European countries. Table 5.1 provides an overview of all company-related variables and Table 5.2 comprises nation-level and economic variables in our dataset, including a detailed definition.

Table 5.1: Definitions, measurements, data sources, and related hypotheses of company-related variables.

Category	Variable	Measurement	Source	Hypothesis
Board variables	Board size	Total number of board members	Datastream	H1
	Board skills	Percentage of members on the board who have either a strong financial background or an industry-specific background	Datastream	H2
	Board structure diversification	Percentage of female board members	Datastream	H3
	Board member affiliations	Average number of other corporate affiliations for the board members	Datastream	H4
	CSR sustainability committee	Dummy variable. Takes value 1 if the company have a CSR committee or team, 0 otherwise	Datastream	H5
Firm variables - CSR reputation	ESG score	Environmental, social, governance performance	Datastream	
Other firm variables	Size	Logarithm of total assets	Datastream	
	Analyst coverage	Total number of analysts providing forecasts to firm's earnings per share	I/B/E/S	
	Cash	The sum of cash and short-term investments divided by total assets	Datastream	
	Leverage	Long-term debt to total assets ratio	Datastream	
	Capex	Capital expenditure divided by total assets times 100	Datastream	
	Earnings variability	Standard deviation of net income before extra items/preferred dividends of the previous five years over total assets	Datastream	
	Price volatility	Average annual stock price movement to a high and low from a mean price for each year	Datastream	
	Return on assets	Earnings before interest, taxes, and depreciation over total assets	Datastream	

Table 5.2: Definitions, measurements, data sources and related hypotheses of nation-level and economic variables.

Category	Variable	Measurement	Source
Political variables	Legislative and corruption	Evaluates government effectiveness, regulatory quality, and absence of corruption	World Bank
	Political participation	Measures voice and accountability	World Bank
	Political stability	Evaluates the likelihood of destabilization or overthrowal of a government	World Bank
Country variables	Herfindahl-Hirschman index (HHI)	Evaluates nation-level market competitiveness	World Bank
	KOF Globalisation index (KOFGI)	Indicates a country's degree of globalization with regard to economic, social, and political dimensions	KOF Swiss Economic Institute
	Growth domestic product (GDP)	Annual growth rate of the gross domestic product	Datastream
	World Press Freedom index (WPI)	Measure for the degree of freedom of the press	RSF Reporters without borders
National culture variables	Power distance index	Extent to which power imbalances are accepted and expected	Hofstede (2001); Hofstede et al. (2010)
	Individualism vs. collectivism	Describes integration of individuals into social groups	Hofstede (2001); Hofstede et al. (2010)
	Uncertainty avoidance index	Describes the level of social tolerance for ambiguous situations	Hofstede (2001); Hofstede et al. (2010)
	Masculinity and femininity	Gender-specific role patterns within a cultural community	Hofstede (2001); Hofstede et al. (2010)

5.3.1.1 Dependent variables to measure corporate controversies

To investigate and measure firm controversies we use the *Refinitiv Controversies score*, which is a scoring methodology that allows for new approaches of academic research by assessing negative news in global media. This score is calculated as an inverse percentile ranking that considers scandals which occur during a company's fiscal year. In addition, this rating is benchmarked on the respective industry group. Its rating methodology consists of 23 concrete ESG controversy topics such as controversies surrounding environmental impact or related to tax fraud (see Refinitiv, 2021).

The *Refinitiv Controversies score* ranges from one to one hundred. The occurrence of scandals has a negative impact on the score of the involved company. Therefore, the more scandals a company exhibits, the lower its score. Furthermore, scandals that entail ongoing legislation disputes as well as lawsuits may also affect the subsequent years and may still be accounted for in controversies ratings of later years. Consequently, this contributes to a distinction in the magnitude of scandals. Companies without any controversies get a score of 100. Refinitiv also already basically takes the market capitalization of the companies and the associated media attention into account.

5.3.1.2 Independent variables

Dorfleitner et al. (2021b) already identify various policy, society, culture, and firm characteristics which influence the occurrence of corporate controversies. Therefore, we use these variables as a starting point for our regressions and add further board level determinants.

Board variables To examine the influence of board structure on scandals, we add further variables which evaluate the board of a company regarding size, the level of expertise, gender diversity, the level of occupancy and distribution of personal resources (time, power, interest), as well as board level CSR efforts.

In detail, *Board size* reveals the total number of board members and is therefore an appropriate quantity to investigate Hypothesis 1. The variable *Board skills* displays the percentage of board members who have either a strong financial or an industry-specific background. Consequently, it is a highly suitable parameter to measure the level of qualification and expertise of a company's board (see Hypothesis 2). In this work, we consider board gender diversity as the proportion of women on board. Therefore, a suitable measure to quantify board gender diversity, as discussed in Hypothesis 3, is the *Board structure diversification* variable, which indicates the percentage of female board members. In order to measure to what extent board members invest personal resources and interest in the respective company and to indicate the level of board busyness as considered by Hypothesis 4, we investigate the *Board member affiliations* variable which measures the average number of other corporate affiliations of board members. To measure the willingness of a firm to extend CSR efforts on the board level, we add the *CSR sustainability committee* dummy variable which takes value 1 if a company has a CSR committee or team and 0 otherwise. It is therefore suitable for the examination of Hypothesis 5.

Firm-related control variables Additionally, we include the variable *ESG score* to quantify a company's overall ESG performance, which is shown to play a major role in the involvement of corporate controversies (see Dorfleitner et al., 2021b). Since the Refinitv controversies score methodology only incorporates three basic firm size categories (large, mid, and small), we use the variables *Size* as well as *Analyst coverage* as a more detailed measuring tool for visibility and media attention. Further firm-related variables to measure capital structure are *Cash* and *Leverage*. The capital expense of a company is measured by the *Capex* variable. Aspects of idiosyncratic firm risk are covered by the variables *Earnings variability* and *Price volatility*. Finally, as a quantity for firm performance, we add the variable *Return on assets (ROA)*.

Political control variables To consider the impact of country-specific political settings and effects we use the *Worldwide governance indicators (WGI)* from World Bank[2]. But as the totality of these governance indicators shows a strong level of collinearity, it cannot simultaneously be used in regressions. The variables *Legislative and corruption*, *Political participation*, and *Political stability* are results of a principal component analysis (PCA) to calculate a set of political factors that can be applied in our statistical analyses. These three variables measure and evaluate aspects of regulatory quality, government effectiveness, rule of law, and control of corruption (*Legislative and corruption*), voice and accountability (*Political participation*), as well as political stability and absence of violence (*Political stability*).

Further country-related control variables Next to political and cultural dimensions, we include the variables *Herfindahl-Hirschman index, KOF Globalisation index, Gross domestic product,* and *World press freedom index* to cover aspects of market competitiveness, the degree of globalization, the overall economic growth, as well as the freedom of the press which potentially affect the occurrence of corporate controversies.

National culture control variables To reflect geographical and country-specific influences on ethical standards and prevailing societal norms as well as corporate behavior, we use the well-known Hofstede cultural dimensions (Hofstede, 2001; Hofstede et al., 2010), namely *Power distance index, Individualism vs. collectivism, Uncertainty avoidance index,* and *Masculinity and femininity.* With the aid of these variables, we are able to examine cross-cultural differences (Beekun and Westerman, 2012) and implement them in our cross-country analyses.

In contrast to all remaining variables in our dataset, these Hofstede's cultural dimensions are time-invariant as they display long-term cultural developments which only fluctuate over generations.

5.3.1.3 Summary Statistics

In order to obtain a large international data universe for further analysis, we include all companies for which all of the observed variables are available. All currency-dependent variables are converted into US dollars. The dataset considers delisted or insolvent firms until the last available rating or financial information to preserve our results from influences of a potential survivorship bias. Table 5.3 shows the descriptive statistics for all variables in our data universe.

[2]see http://info.worldbank.org/governance/wgi/

Table 5.3: Descriptive statistics.

Statistic	Mean	St. Dev.	Min	Median	Max
Controversies score	90.84	22.71	1.00	100.00	100.00
Board size	10.06	3.29	1.00	10.00	36.00
Board skills	55.36	23.39	0.00	55.56	100.00
Board structure diversification	14.03	12.05	0.00	12.50	100.00
Board member affiliations	1.09	0.89	0.00	0.91	14.75
CSR sustainability committee	0.49	0.50	0.00	0.00	1.00
ESG score	42.87	20.16	0.11	40.97	95.21
Size	15.33	1.52	8.68	15.31	21.41
Analyst coverage	12.62	8.12	1.00	11.00	56.00
Cash	0.14	0.15	0.00	0.09	1.03
Leverage	0.22	0.18	0.00	0.20	3.88
Capex	5.82	7.40	−6.41	3.90	226.60
Earnings variability	0.05	0.13	0.00	0.02	7.19
Price volatility	28.39	9.84	5.90	26.79	81.13
Return on assets	0.11	0.18	−14.84	0.11	8.90
Legislative and corruption	0.79	0.71	−2.56	0.96	2.57
Political participation	0.38	0.64	−2.87	0.52	1.26
Political stability	−0.10	0.57	−3.23	−0.06	1.29
Herfindahl-Hirschman index	0.10	0.12	0.03	0.06	0.71
KOF Globalisation index	80.04	7.00	58.00	81.51	90.98
Gross domestic product	2.26	2.18	−9.13	2.30	25.16
World press freedom index	19.35	16.90	−10.00	17.00	136.00
Power distance index	46.64	15.28	11.00	40.00	104.00
Individualism vs. collectivism	72.11	23.97	13.00	89.00	91.00
Uncertainty avoidance index	52.60	19.03	8.00	46.00	112.00
Masculinity and femininity	60.52	15.75	5.00	62.00	95.00

This table presents the mean, standard deviation, median, minimum, and maximum values of all variables of the full dataset ($N = 38,997$). All variables are as described in Table 5.1 and Table 5.2.

5.3.2 Methodology

Our data is subject to different structures and frequencies. While the board structure variables, as well as firm-, political-, and country-related variables are calculated on a yearly basis, Hofstede's cultural dimensions are time-invariant. We examine the variance inflation factor (VIF) values based on an OLS regression model, to verify for potential multicollinearity between variables in our dataset. The outcome indicates no linear relations for any of our variables.

Following Dorfleitner et al. (2021b), we focus on considering variation over time and between the companies. To evaluate both, within and between effects, simultaneously in one model, our calculations rely on a hybrid regression model (see Allison, 2009; Schunck, 2013). This hybrid model is basically defined by

$$y_{it} = \beta_0 + \beta_1(x_{it} - \bar{x}_i) + \beta_2 c_i + \beta_3 \bar{x}_i + \mu_i + \epsilon_{it} \tag{5.1}$$

where y_{it} represents the dependent variable for an individual i at time t. Moreover, x_{it} denotes a variable that varies over both, time and individuals, whereas the variable c_i varies only over individuals. Furthermore, $\bar{x}_i$ denotes the mean of the x_{it} for a fixed i over t. Aside from that, μ_i is an error term and random intercept, while ϵ_{it} describes a noise variable. By using (5.1) we are able to investigate both the within effect (β_1) and the between effect (β_3) in single models, while keeping time-invariant effects (β_2).

5.4 Results

The results of the hybrid regression model are presented in Table 5.4. First, we analyze the contemporary[3] controversies score as a dependent variable. The results comprise the within and between effects. In that respect, the within results estimate the effects of changes over time, whereas the between results of the hybrid regression examine the cross-section of firms.

When considering the variable *Board size*, we observe a positive and significant coefficient in the within part as well as a negative and significant effect in the between part of the hybrid regression. This indicates that an increase of *Board size* over time leads to an increase of the *Controversies score* whereas companies with larger boards tend to have a lower *Controversies score* than companies with fewer board members.

Coles et al. (2008) publish a negative relation between board size and firm risk. There-

[3]This means that the independent variables are not lagged.

Table 5.4: Hybrid regression - results of the full dataset.

		Within effects		Between effects	
Category	Variable	Coefficient	S.E.	Coefficient	S.E.
Board structure	Board size	0.1982***	0.0737	−0.1526**	0.0702
	Board skills	0.0291***	0.0067	0.0447***	0.0105
	Board structure diversification	0.0187	0.0160	−0.0384**	0.0192
	Board member affiliations	0.1633	0.1957	−0.8357***	0.2538
	CSR sustainability committee	−1.3137***	0.3849	−0.2772	0.5237
Firm variables	ESG score	−0.0764***	0.0131	−0.1168***	0.0140
	Size	−2.0915***	0.3421	−3.9255***	0.1816
	Analyst coverage	−0.0485	0.0321	−0.1854***	0.0299
	Cash	−0.6299	1.6141	−5.8535***	1.3870
	Leverage	1.0379	1.1894	4.3059***	1.0638
	Capex	0.0567***	0.0219	−0.0344	0.0283
	Earnings variability	−1.1984	1.4694	−2.9677**	1.4794
	Price volatility	−0.2235***	0.0286	−0.2151***	0.0218
	Return on assets	2.5591***	0.7522	−1.7072	1.3776
Political variables	Legislative and corruption	4.3966***	1.3025	2.6284***	0.5701
	Political participation	1.6999	1.4984	2.6473***	0.8583
	Political stability	0.2592	0.9034	1.9826***	0.4982
Country variables	Herfindahl-Hirschman index	−1.8040	8.8855	−0.6320	1.5794
	KOF Globalisation index	−0.4786***	0.1348	0.0343	0.0477
	Gross domestic product	−0.0842	0.0606	0.0161	0.2027
	World press freedom index	0.0695***	0.0160	0.1576***	0.0399
National culture variables	PDI			0.0510**	0.0233
	IDV			−0.1312***	0.0161
	MAS			−0.0587***	0.0149
	UAI			0.0461***	0.0146
	Pseudo R^2 (total)	0.32			

This table shows the results derived from the within-between regression based on the full sample. Coefficients of within-regression (β_1) and between-regression (β_2 and β_3) results, standard errors, and pseudo R^2 are reported upon. All variables are as described in Table 5.1 and Table 5.2. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

fore, we attribute the observed within coefficient to a more future-oriented and farsighted decision-making process, which also leads to a decrease in risk to be involved in corporate controversies. Moreover, one possible explanation for the negative between effect could be the following. Since controversies could generate a high level of attention very quickly, fast and efficient crisis management by the company concerned is important. In line with prior studies (see De Andres et al., 2005; Lipton and Lorsch, 1992; Jensen, 1993), large boards also face disadvantages of larger groups regarding poor coordination, less flexibility, as well as communication issues which may all decrease board efficiency and lead to larger impacts of corporate controversies.

Overall, we find mixed evidence regarding *Board size* and, therefore, at this stage, we can not confirm Hypothesis 1.

The *Board skills* variable exhibits a positive and strongly significant coefficient in the within and between part of the hybrid regression. Both results are significant at the 1% level, which illustrates that there is a positive relationship between *Board skills* and the

Controversies score.

One explanation for this observation might be intuition since board members with strong financial or industry-specific backgrounds may assess more fully the risks of behavior that potentially leads to corporate controversies. This would correspond to the findings of Coles et al. (2008) which indicate a negative relation between board size and firm risk.

Furthermore, the results also indicate that firms with higher-skilled boards tend to have significantly better *Controversies scores* than competitors with lower-skilled boards. One possible reason for this finding may be that companies with less skilled board members may consider unethical practices to accomplish competitive advantages or to avoid competitive disadvantages.

In sum, we find evidence that companies with high values of *Board skills* are less likely to be involved in corporate controversies. Therefore, our results support Hypothesis 2.

To examine Hypothesis 3, we investigate the *Board structure diversification* indicating the percentage of female board members. The coefficient of the within effect is positive but insignificant, while the between effect reveals a negative and significant value. Previous literature finds a negative association of higher gender diversity and firm-level corporate social irresponsibility (see Jain and Zaman, 2020), which also tends to be evident from the within part of the results. However, when comparing between firms, the regression indicates a negative value. One possible explanation may be the following. As already published by McCabe et al. (2006), there are no differences in the overall perceptions of ethical behaviors between sexes, but there are differences in the perception of strength of unethical actions (i.e., bribery), namely women perceive bribery as significantly less ethical than males. Furthermore, Valentine and Rittenburg (2007) illustrate that females tend to exhibit greater intentions to act more ethically.

As a result, increasing gender diversification may increase monitoring efforts. Following Adams and Ferreira (2009), tougher monitoring may lead managers to be less willing to share information. This entrenchment effect may therefore achieve exactly the opposite effect as unsupervised roguish managers may lead to an increase of corporate irresponsible behavior.

Besides, another aspect cannot be ignored. Since only very few observations in our dataset exhibit a *Board structure diversification* value above 50, we classify the board of directors as a male-dominated profession (in the style of Cumming et al., 2015). Therefore, women board members are often outnumbered and thus may face difficulties in asserting themselves against their male counterparts on board as well as management levels.

Thus, we attribute the observed effect not only to the discrepancy in the perception of unethical actions which in turn leads to the detection and open communication of questionable behavior but also to rather fewer chances of assertiveness of female board

members against entrenchment effects of roguish managers.

All in all, we could not confirm Hypothesis 3 and even find first evidence of a rather opposing relationship.

Considering the *Board member affiliations* variable, we observe evidence in favor of Hypothesis 4, since the coefficients of the between effect show a negative and significant value on the 1%-level. Therefore, an increasing number of other corporate affiliations of board members leads to a decrease in the *Controversies score*.

One possible explanation may be that board members with high numbers of corporate affiliations have difficulty following each of their posts with maximum attention and commitment. Therefore, companies with high *Board member affiliations* values are more likely to become involved in scandals compared to competitors with less numerous *Board member affiliations*.

Whilst considering the *CSR sustainability committee* as a measure for the willingness to extend CSR efforts on the board level, we detect the following results. The coefficient of the *CSR sustainability committee* variable reveals a negative and significant value on 1% level, which indicates that there is a negative relationship between *CSR sustainability committee* and the *Controversies score* as well as a negative but insignificant between effect. In line with Dorfleitner et al. (2021b), one possible explanation may be that companies tend to rely on an insurance-like effect of high levels of CSR values and therefore the incentive for ethical behavior also decreases. Even Dorfleitner et al. (2020) find evidence for a similar relation when examining the relationship between corporate social performance and corporate financial performance. In summary, we find evidence that despite a CSR committee companies with good CSR reputations are more likely to be involved in corporate controversies, which indicates evidence in favor of Hypothesis 5.

Regarding firm, political, country, and cultural variables, our findings are in line with Dorfleitner et al. (2021b) but shall not be elaborated in detail in this paper. To put the most important findings in a nutshell, on the one hand, companies with good CSR ratings (*ESG score*), good capital structure (high *Cash* and low *Leverage*), as well as riskier (*Price volatility* and *Earnings variability*), larger (*Size*), and high-attention (*Analyst coverage*) companies as well as companies located in countries with high *IDV* and *MAS* values exhibit lower *Controversies score* values and are therefore more likely to become involved in corporate controversies.

On the other hand, companies which are located in countries with strong political conditions (*Legislative and corruption*, *Political participation* and *Political stability*) as well as in cultures with high *Power distance index* and *Uncertainty avoidance index* values are less likely to become involved in a corporate controversy.

5.5 Robustness checks

To judge the robustness of our results, we run some further regressions. First, to test whether some of our variables reveal evidence for u-shaped relations, we add quadratic predictor variables for *Board size, Board skills, Board structure diversification*, and *Board member affiliations* and run the hybrid regression again. Since none of these variables show significant coefficients in both, i.e., predictor and quadratic predictor variables, we find no evidence for u-shaped relations. As there are no new insights from these results, we do not report them in this paper.

As another robustness test, we divide our data sample into a small boards sample (< 10 members) and a large boards sample (≥ 10 members) and run the hybrid regression again to investigate whether some of the observed effects rely on board size. The results are displayed in Table 5.5 and Table 5.6. Regarding the *Board size* variable we detect a positive and significant coefficients from the within results of the large board sample while the within coefficient of the small board sample is positive but insignificant. Both between effects are insignificant, what we contribute to rather small differences between the entities in the respective samples. Thus, we attribute these results in favor of Hypothesis 1, even if we restrict them to larger boards. Taking the *Board skills* variable into account, the within and between coefficient in the large board sample and the between effect from the small board sample are positive and significant. Thus, we again detect supporting evidence in favor of Hypothesis 2. When considering the *Board structure diversification*, we observe a negative and significant coefficient in the large board sample and no significant coefficients in the small board sample. Therefore, we still reject Hypothesis 3. The *Board member affiliations* reveals a strongly significant and negative coefficient of the between effect in both, large and small board samples. Surprisingly, the coefficient of the within effect of the large board sample is positive and significant on 10% level, whereas the respective coefficient of the small board sample is negative and insignificant. One possible explanation may be that large boards are able to compensate for the effect of decreasing attention and commitment. Thus, this observed effect is particularly strong for small boards. All in all, we continue to support Hypothesis 4 but add that this effect applies especially to small boards. Regarding *CSR sustainability committee*, the results in both samples are in line with previous results and therefore we again find supporting evidence that companies with high CSR efforts are more likely to become involved in corporate controversies.

Table 5.5: Hybrid regression - results of the large board dataset.

		Within effects		Between effects	
Category	Variable	Coefficient	S.E.	Coefficient	S.E.
Board structure	Board size	0.2913**	0.1172	−0.0918	0.1336
	Board skills	0.0451***	0.0108	0.0491***	0.0175
	Board structure diversification	0.0308	0.0281	−0.0683**	0.0343
	Board member affiliations	0.5987*	0.3095	−1.0173**	0.4173
	CSR sustainability committee	−1.2982**	0.5950	−0.2951	0.8366
Firm variables	ESG score	−0.0833***	0.0203	−0.1248***	0.0218
	Size	−2.7363***	0.5916	−5.4801***	0.2990
	Analyst coverage	−0.0120	0.0487	−0.1538***	0.0457
	Cash	−1.0704	2.9321	−7.7687***	2.6860
	Leverage	2.8910	2.0691	5.1019***	1.8286
	Capex	0.0633	0.0483	−0.1476***	0.0565
	Earnings variability	−3.2491	3.3176	−7.4071*	4.3043
	Price volatility	−0.2089***	0.0478	−0.2492***	0.0382
	Return on assets	4.6024***	1.5542	−5.4970	3.4523
Political variables	Legislative and corruption	3.4705*	2.0250	3.8902***	0.8418
	Political participation	−0.2092	2.2686	4.0772***	1.2419
	Political stability	−0.3328	1.3981	3.1650***	0.7630
Country variables	Herfindahl-Hirschman index	2.9275	14.1402	−0.9549	2.5503
	KOF Globalisation index	−0.5283***	0.1984	−0.0876	0.0775
	Gross domestic product	−0.0055	0.0908	0.2057	0.2902
	World press freedom index	0.0918***	0.0234	0.2011***	0.0553
National culture variables	PDI			0.0566	0.0353
	IDV			−0.1109***	0.0261
	MAS			−0.1020***	0.0242
	UAI			0.0730***	0.0224
	Pseudo R^2 (total)	0.36			

This table shows the results derived from the within-between regression based on the large board sample. Coefficients of within-regression (β_1) and between-regression (β_2 and β_3) results, standard errors, and pseudo R^2 are reported upon. All variables are as described in Table 5.1 and Table 5.2. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 5.6: Hybrid regression - results of the small board dataset.

		Within effects		Between effects	
Category	Variable	Coefficient	S.E.	Coefficient	S.E.
Board structure	Board size	0.1971	0.1726	0.0714	0.1651
	Board skills	0.0046	0.0088	0.0395***	0.0105
	Board structure diversification	0.0060	0.0198	−0.0222	0.0191
	Board member affiliations	−0.1139	0.2717	−0.7953***	0.2584
	CSR sustainability committee	−1.8247***	0.5249	−0.3178	0.5448
Firm variables	ESG score	−0.0524***	0.0184	−0.0698***	0.0153
	Size	−1.4091***	0.4443	−2.4725***	0.2049
	Analyst coverage	−0.1160**	0.0458	−0.2103***	0.0334
	Cash	−1.1542	1.8982	−2.5424*	1.3597
	Leverage	−0.1400	1.4482	2.4146**	1.0851
	Capex	0.0520**	0.0229	0.0596**	0.0272
	Earnings variability	−0.0118	1.5630	−1.4764	1.3290
	Price volatility	−0.2060***	0.0376	−0.1842***	0.0222
	Return on assets	1.1380	0.8328	−0.3397	1.1633
	Legislative and corruption	4.1245**	1.8518	1.8363***	0.6515
	Political participation	3.7595*	2.1740	1.1058	0.9567
	Political stability	0.4715	1.2695	1.2841**	0.5499
	Herfindahl-Hirschman index	−16.0696	11.9744	0.4926	1.7318
	KOF Globalisation index	−0.4815**	0.2045	0.0615	0.0514
	Gross domestic product	−0.2243***	0.0847	0.0530	0.2002
	World press freedom index	0.0522**	0.0232	0.0999**	0.0403
National culture variables	PDI			0.0244	0.0289
	IDV			−0.1009***	0.0182
	MAS			−0.0357**	0.0162
	UAI			0.0350**	0.0174
	Pseudo R^2 (total)	0.24			

This table shows the results derived from the within-between regression based on the small board sample. Coefficients of within-regression (β_1) and between-regression (β_2 and β_3) results, standard errors, and pseudo R^2 are reported upon. All variables are as described in Table 5.1 and Table 5.2. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

5.6 Conclusion

In this paper, we examine an international dataset that comprises over 6,100 companies located in 44 countries and associated measurements for social irresponsible behavior, board structures, as well as further firm-related, political, nation-level economical, and cultural variables from 2002 to 2018.

We investigate the impact of various board structure parameters, namely size, qualification and skills, gender diversity, busyness, and ESG efforts on the occurrence of corporate controversies. Our calculations are based on the within-between hybrid regression model. The results show that we can identify board variables that tend to increase as well as variables that tend to reduce a firm's irresponsible corporate behavior. On the one hand, firms associated with high CSR efforts (measured by the existence of a *CSR sustainability committee*) as well as high levels of *Board member affiliations* are more likely to become involved in a corporate controversy. On the other hand, companies with high levels of *Board skills* are less likely to be involved in a corporate controversy. Even companies with larger boards tend to be less likely involved in a corporate controversies.

No clear evidence could be determined from *Board structure diversification*.

One potential limitation of this study lies in the fact that controversies scores as well as most of our variables are only calculated once per year. Future research could focus on more detailed short-, medium- and long-term effects as well as on a closer investigation of the influence of board characteristics regarding specific countries and cultures.

Our work provides enhanced aspects towards the influence of board structure on the occurrence of corporate controversies. Additionally, we implement new approaches and ideas to extend existing literature of corporate social irresponsibility, still holding promising potential for further research.

Chapter 6

Conclusion

Besides its contribution to academic literature, this dissertation reveals several implications for investors, managers, and policy makers.

Investor implications In particular, the topic of SRI has been receiving increasing attention from private and institutional investors. One major field of interest regarding the CSP–CFP relation deals with the question of whether or not there are significant performance advantages or disadvantages of SRI strategies. The first research paper in this thesis sheds further light on this issue and observes that regardless of stock-weighting strategies almost all of the considered ESG portfolios which consist of top-rated companies indicate no significant underperformance. These observations are good news for value-driven investors. By introducing corporate controversies as an additional dimension of ESG, this work offers a practical and easy-to-replicate way for investors to consider not only the level of "good" activities of companies but also to evaluate whether "bad" practices, in terms of social irresponsible or unethical behavior, are avoided. In the context of portfolio selection, the results of this thesis show that investors should focus mainly on smaller companies with low ESG ratings as well as clean-coated firms, i.e., firms with no controversies, in order to achieve abnormal returns. For those who attach importance to ESG and controversies ratings, this provides a vast opportunity to reward better scoring placements of companies and also strategies to gain higher returns.

Apart from implications for stock-pickers, there are also noteworthy results for investors, who prefer to choose SR mutual funds. Since the short and long-term CSI and ESG persistence of SR mutual funds are preserved, there is, generally speaking, no need for ethically-motivated passive investors to permanently monitor their SR mutual fund portfolio.

Moreover, value-driven and ethically-motivated investors who seek to invest their capital

into companies which act ethically cannot simply choose SR funds with a high ESG rating. Since SR mutual funds with a very high ESG rating tend to have low controversies ratings and vice versa, ethical investors have to verify funds in which they want to invest very closely and may have to choose one side in their investment decisions. Moreover, the data provider Refinitiv Lipper's fund database has recently introduced ESG as well as ESG controversies ratings for several funds. Hence, investors may not have to calculate ESG and ESG controversies ratings themselves. Regarding the ESG performance of SR mutual funds our results indicate that high-paid fund managers do not guarantee higher fund ratings, as ESG-based social performance of high-paid managers is clearly worse than that of the lower-paid managers. At least high-paid managers surpass their lower-paid colleagues regarding controversies-based performance.

With respect to the occurrence of corporate scandals, this dissertation offers indications that cultural, political, societal, as well as firm-related characteristics increase or reduce the occurrence of corporate controversies. Many factors that influence patterns of unethical behavior are intuitively appealing and can easily be adopted by investors. In this regard, this work may provide investors with an initial indication under what circumstances companies are more likely or less likely to be involved in a corporate scandal. Consequently, the results of this thesis can immediately be employed for the implementation into stock selecting processes for both private and professional investors.

Managerial implications The managerial implications of this work are quite straightforward. First, fund managers, in particular those who manage SR mutual funds, may use an additional evaluation concerning CSI criteria in their portfolio-building process. Particularly for ethically motivated investors, good CSI performance and the avoidance of unethical corporate behavior may play a major role in their investment decisions. Additionally, the results indicate evidence that the CSR performance of high-paid SR fund managers could be improved, but do not advise fund managers to exclusively focus on CSR or CSI criteria.

Second, managers, as well as executives, pay great attention to and put tremendous effort into implementing CSR activities in order to obtain better CSP for their company. It is also crucial to consider CSI as the "dark side of CSP" and to understand determinants that affect the CSI performance of their company. This work gains an insight into cross-national and intra-industrial differences between patterns of unethical corporate behavior and thus helps managers to comprehend not only firm-specific key factors but also those factors outside their own company and, therefore, also beyond their control.

Third, the results are of special benefit to ethically motivated managers who may find implications for their evaluation as well as decision-making processes. They may help to avoid and uncover patterns of entrenchment as well as a failure of monitoring.

Policy implications By identifying nation-level variables that influence the occurrence of unethical corporate behavior, this work is also of particular interest for policy makers as it offers important policy implications. By providing empirical evidence, it contributes to sensitize policy makers towards the role of political institutions within the context of corporate controversies. In more detail, the results may help policy makers in emerging and less developed countries, in which rules or policies of politics, society, as well as financial systems are still evolving or need to be redefined. Moreover, it may provide reasons for (re)consideration in the term of unethical behavior when defining sustainability and social responsibility norms and goals. Policy makers in developed as well as emerging countries should adopt terms of unethical behaviors besides ESG criteria for better decisions in rating and awarding processes, such as public procurement or even governmental aid for companies in precarious financial situations. With regard to SR mutual funds and in order to obtain the "socially responsible" label, it is worth considering incorporating additional criteria and hurdles with regard to unethical actions by companies in their holdings. This idea goes beyond the often practiced negative screening and requires controversies as an additional ESG dimension, which may ensure that "fund-greenwashing" is avoided and investors are not misled.

Limitation and further research Since the calculations of all research papers in this dissertation are largely or wholly based on the Refintiv ESG controversies score and ESG score, potential limitations lie in restrictions of these scoring metrics. More precisely, both scores are only calculated once per year. Thus, especially in the case of corporate scandals in a portfolio-building context (Dorfleitner et al., 2020), future research should implement a higher evaluation frequency as well as a faster incorporation of the occurrence of corporate controversies into their ratings. This may allow a close and more detailed investigation of short-, mid-, and long-term effects as well as stock market reactions. Since even other data providers, such as MSCI or ISS (Institutional Shareholder Services group of companies), offer firm ratings based on ESG controversies, a comparison of stock portfolios founded in different ratings and cut-offs may be fruitful for researchers.

Future work that investigates corporate controversies on firm-level may also focus on the influence of industry sectors on the occurrence of corporate controversies as well as it may highlight industry-specific differences. To provide profound causal evidence for single drivers of controversies, deeper insights into the motivations and backgrounds of the actors involved may be necessary. In this regard, experimental as well as interview-based research appears to be promising (Dorfleitner et al., 2021b). Besides that, one might shed further light on the key role of executives in the occurrence of corporate controversies. Characteristics such as age, gender, cultural background might be worth a closer examination. Next to the examination of an aggregated ESG controversies score or the total

sum of controversies, further research may focus on specific controversy topics (such as business ethics controversies, human rights controversies, or environmental controversies). The effects and influences of scandals in various categories have been not yet sufficiently investigated and hold promising potential for a range of further research. Possible research questions in this respect are: How do markets react to different topics of controversies? Are there national or cultural differences in the occurrence of scandals regarding different controversies topics? How do corporate controversies of various topics differ in terms of medium- and long-term effects on companies?

In the context of mutual funds, some of the results in early years are driven by only few observations due to data limitations (Dorfleitner et al., 2021a). However, this work is intended to provide a proper first step toward examining CSI in the context of SR mutual funds and allows scope for further research based on larger amounts of data. The methodology to examine the CSI performance of mutual funds provides a range of possibilities for further research. On the one hand, future research may seek to comprehend the CSI performance of conventional and SR mutual funds and examine whether SR mutual funds perform better on CSI than conventional funds. On the other hand, another interesting research question is whether, and if so to what extent, SR and conventional mutual funds with high and low CSI ratings differ in terms of risk. Apart from that, fund managers' characteristics such as age, years of education, experience in relation to their funds' CSR or CSI performance might be worth a closer investigation. Last but not least, an examination of potential out- or underperformance of funds with high and low CSI may also hold promising potential.

Bibliography

Adams, R. B., Ferreira, D., 2007. A theory of friendly boards. The Journal of Finance 62 (1), 217–250.

Adams, R. B., Ferreira, D., 2009. Women in the boardroom and their impact on governance and performance. Journal of Financial Economics 94 (2), 291–309.

Aguilera, R. V., Rupp, D. E., Williams, C. A., Ganapathi, J., 2007. Putting the S back in corporate social responsibility: A multilevel theory of social change in organizations. Academy of Management Review 32 (3), 836–863.

Ahmed, K., Hossain, M., Adams, M. B., 2006. The effects of board composition and board size on the informativeness of annual accounting earnings. Corporate Governance: An International Review 14 (5), 418–431.

Alas, R., 2006. Ethics in countries with different cultural dimensions. Journal of Business Ethics 69 (3), 237–247.

Aldrich, H. E., Fiol, C. M., 1994. Fools rush in? The institutional context of industry creation. Academy of Management Review 19 (4), 645–670.

Allison, P. D., 2009. Fixed effects regression models. Vol. 160. SAGE Publications, Newbury Park, CA.

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

Argandoña, A., 2001. Corruption: The corporate perspective. Business Ethics: A European Review 10 (2), 163–175.

Arora, P., Dharwadkar, R., 2011. Corporate governance and corporate social responsibility (CSR): The moderating roles of attainment discrepancy and organization slack. Corporate Governance: An International Review 19 (2), 136–152.

Ashforth, B. E., Gibbs, B. W., 1990. The double-edge of organizational legitimation. Organization Science 1 (2), 177–194.

Ashforth, B. E., Mael, F., 1989. Social identity theory and the organization. Academy of Management Review 14 (1), 20–39.

Auer, B. R., 2016. Do socially responsible investment policies add or destroy European stock portfolio value? Journal of Business Ethics 135 (2), 381–397.

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.

Aupperle, K. E., Carroll, A. B., Hatfield, J. D., 1985. An empirical examination of the relationship between corporate social responsibility and profitability. Academy of Management Journal 28 (2), 446–463.

Baldini, M., Dal Maso, L., Liberatore, G., Mazzi, F., Terzani, S., 2018. Role of country- and firm-level determinants in environmental, social, and governance disclosure. Journal of Business Ethics 150 (1), 79–98.

Baraibar-Diez, E., D Odriozola, M., 2019. CSR committees and their effect on ESG performance in UK, France, Germany, and Spain. Sustainability 11 (18), 5077.

Barnea, A., Rubin, A., 2010. Corporate social responsibility as a conflict between shareholders. Journal of Business Ethics 97 (1), 71–86.

Barnett, M. L., 2007. Stakeholder influence capacity and the variability of financial returns to corporate social responsibility. Academy of Management Review 32 (3), 794–816.

Barreda-Tarrazona, I., Matallín-Sáez, J. C., Balaguer-Franch, M. R., 2011. Measuring investors' socially responsible preferences in mutual funds. Journal of Business Ethics 103 (2), 305–330.

Bathula, H., 2008. Board characteristics and firm performance: Evidence from New Zealand. Ph.D. thesis, Auckland University of Technology.

Bauer, R., Koedijk, K., Otten, R., 2005. International evidence on ethical mutual fund performance and investment style. Journal of Banking & Finance 29 (7), 1751–1767.

Bauer, T. N., Aiman-Smith, L., 1996. Green career choices: The influence of ecological stance on recruiting. Journal of Business and Psychology 10 (4), 445–458.

Becker, G. S., 1968. Crime and punishment: An economic approach. Journal of Political Economy 76 (2), 169–217.

Beekun, R. I., Westerman, J. W., 2012. Spirituality and national culture as antecedents to ethical decision-making: A comparison between the United States and Norway. Journal of Business Ethics 110 (1), 33–44.

Bello, Z. Y., 2005. Socially responsible investing and portfolio diversification. Journal of Financial Research 28 (1), 41–57.

Bhattacharya, C. B., Sen, S., 2004. Doing better at doing good: When, why, and how consumers respond to corporate social initiatives. California Management Review 47 (1), 9–24.

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

Boulouta, I., 2013. Hidden connections: The link between board gender diversity and corporate social performance. Journal of Business Ethics 113 (2), 185–197.

Bowen, R. M., Chen, X., Cheng, Q., 2008. Analyst coverage and the cost of raising equity capital: Evidence from underpricing of seasoned equity offerings. Contemporary Accounting Research 25 (3), 657–700.

Boyle, E. J., Higgins, M. M., Rhee, G. S., 1997. Stock market reaction to ethical initiatives of defense contractors: Theory and evidence. Critical Perspectives on Accounting 8 (6), 541–561.

Brammer, S., Millington, A., Rayton, B., 2007. The contribution of corporate social responsibility to organizational commitment. The International Journal of Human Resource Management 18 (10), 1701–1719.

Breusch, T. S., 1978. Testing for autocorrelation in dynamic linear models. Australian Economic Papers 17 (31), 334–355.

Breusch, T. S., Pagan, A. R., 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47 (5), 1287–1294.

Brown, T. J., Dacin, P. A., 1997. The company and the product: Corporate associations and consumer product responses. Journal of Marketing 61 (1), 68–84.

Brundtland, G. H., 1987. Report of the world commission on environment and development: Our common future. United Nations General Assembly Document A/42/427, Oslo.

Business Ethics Quarterly, 2002. Appendix: Principles of stakeholder management. Business Ethics Quarterly 12 (2), 257–264.

Cai, Y., Pan, C. H., Statman, M., 2016. Why do countries matter so much in corporate social performance? Journal of Corporate Finance 41, 591–609.

Campbell, J. L., 2007. Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. Academy of Management Review 32 (3), 946–967.

Capelle-Blancard, G., Monjon, S., 2012. Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. Business Ethics: A European Review 21 (3), 239–250.

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.

Carter, M. E., Franco, F., Gine, M., 2017. Executive gender pay gaps: The roles of female risk aversion and board representation. Contemporary Accounting Research 34 (2), 1232–1264.

Chapple, L., Humphrey, J. E., 2014. Does board gender diversity have a financial impact? Evidence using stock portfolio performance. Journal of Business Ethics 122 (4), 709–723.

Chatterji, A. K., Durand, R., Levine, D. I., Touboul, S., 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. Strategic Management Journal 37 (8), 1597–1614.

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

Chevalier, J., Ellison, G., 1999. Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance. The Journal of Finance 54 (3), 875–899.

Choi, B. B., Lee, D., Park, Y., 2013. Corporate social responsibility, corporate governance and earnings quality: Evidence from Korea. Corporate Governance: An International Review 21 (5), 447–467.

Claringbould, I., Knoppers, A., 2007. Finding a 'normal' woman: Selection processes for board membership. Sex Roles 56 (7), 495–507.

Coffey, B. S., Wang, J., 1998. Board diversity and managerial control as predictors of corporate social performance. Journal of Business Ethics 17 (14), 1595–1603.

Coles, J. L., Daniel, N. D., Naveen, L., 2008. Boards: Does one size fit all? Journal of Financial Economics 87 (2), 329–356.

Coluccia, D., Fontana, S., Solimene, S., 2018. Does institutional context affect CSR disclosure? A study on Eurostoxx 50. Sustainability 10 (8), 2823.

Cortez, M. C., Silva, F., Areal, N., 2009. The performance of European socially responsible funds. Journal of Business Ethics 87 (4), 573–588.

Crilly, D., Zollo, M., Hansen, M. T., 2012. Faking it or muddling through? Understanding decoupling in response to stakeholder pressures. Academy of Management Journal 55 (6), 1429–1448.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. Journal of Economic Literature 47 (2), 448–74.

Crossland, C., Hambrick, D. C., 2011. Differences in managerial discretion across countries: How nation-level institutions affect the degree to which CEOs matter. Strategic Management Journal 32 (8), 797–819.

Cumming, D., Leung, T. Y., Rui, O., 2015. Gender diversity and securities fraud. Academy of Management Journal 58 (5), 1572–1593.

Cuthbertson, K., Nitzsche, D., O'Sullivan, N., 2008. UK mutual fund performance: Skill or luck? Journal of Empirical Finance 15 (4), 613–634.

Darmadi, S., 2013. Board members' education and firm performance: Evidence from a developing economy. International Journal of Commerce and Management 23 (2), 113–135.

De Andres, P., Azofra, V., Lopez, F., 2005. Corporate boards in OECD countries: Size, composition, functioning and effectiveness. Corporate Governance: An International Review 13 (2), 197–210.

Derwall, J., Guenster, N., Bauer, R., Koedijk, K., 2005. The eco-efficiency premium puzzle. Financial Analysts Journal 61 (2), 51–63.

Devinney, T. M., Schwalbach, J., Williams, C. A., 2013. Corporate social responsibility and corporate governance: Comparative perspectives. Corporate Governance: An International Review 21 (5), 413–419.

Dhaliwal, D. S., Li, O. Z., Tsang, A., Yang, Y. G., 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. The Accounting Review 86 (1), 59–100.

Dharwadkar, R., Guo, J., Shi, L., Yang, R., 2021. Corporate social irresponsibility and boards: The implications of legal expertise. Journal of Business Research 125, 143–154.

Dickson, M., BeShears, R., Gupta, V., 2004. The impact of societal culture and industry on organizational culture: Theoretical explanations. In: House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., Gupta, V. (Eds.), Culture, leadership, and organizations: The GLOBE study of 62 societies. SAGE Publications, Thousand Oaks, CA, pp. 4–93.

DiMaggio, P. J., Powell, W. W., 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review 48 (2), 147–160.

Dorfleitner, G., Halbritter, G., Nguyen, M., 2015. Measuring the level and risk of corporate responsibility–An empirical comparison of different ESG rating approaches. Journal of Asset Management 16 (7), 450–466.

Dorfleitner, G., Kreuzer, C., Laschinger, R., 2021a. How socially irresponsible are socially responsible mutual funds? A persistence analysis. Finance Research Letters 43, 101990.

Dorfleitner, G., Kreuzer, C., Sparrer, C., 2020. ESG controversies and controversial ESG: About silent saints and small sinners. Journal of Asset Management 21 (5), 1–20.

Dorfleitner, G., Kreuzer, C., Sparrer, C., 2021b. It's not only size that matters: On the influence of policy, society, culture, and firm characteristics on corporate controversies. SSRN Working Paper 3956427.

Dorfleitner, G., Leidl, M., Reeder, J., 2012. Theory of social returns in portfolio choice with application to microfinance. Journal of Asset Management 13 (6), 384–400.

Dorfleitner, G., Utz, S., Wimmer, M., 2018. Patience pays off–corporate social responsibility and long-term stock returns. Journal of Sustainable Finance & Investment 8 (2), 132–157.

Dowling, J., Pfeffer, J., 1975. Organizational legitimacy: Social values and organizational behavior. Pacific Sociological Review 18 (1), 122–136.

Drempetic, S., Klein, C., Zwergel, B., 2019. The influence of firm size on the ESG score: Corporate sustainability ratings under review. Journal of Business Ethics 167 (2), 333–360.

Durand, R., Jacqueminet, A., 2015. Peer conformity, attention, and heterogeneous implementation of practices in MNEs. Journal of International Business Studies 46 (8), 917–937.

Durbin, J., Watson, G. S., 1971. Testing for serial correlation in least squares regression. III. Biometrika 58 (1), 1–19.

Eberhardt-Toth, E., 2017. Who should be on a board corporate social responsibility committee? Journal of Cleaner Production 140, 1926–1935.

Edmans, A., 2011. Does the stock market fully value intangibles? Employee satisfaction and equity prices. Journal of Financial Economics 101 (3), 621–640.

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

Elkington, J., 1997. Partnerships from cannibals with forks: The triple bottom line of 21st-century business. Capstone, Oxford.

Ertimur, Y., Ferri, F., Muslu, V., 2011. Shareholder activism and CEO pay. The Review of Financial Studies 24 (2), 535–592.

Fama, E., 1965. The behavior of stock-market prices. The Journal of Business 38 (1), 34–105.

Fama, E., 1970. Efficient capital markets: A review of theory and empirical work. The Journal of Finance 25 (2), 383–417.

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

Fama, E. F., Jensen, M. C., 1983. Separation of ownership and control. The Journal of Law and Economics 26 (2), 301–325.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81 (3), 607–636.

Fauser, D. V., Utz, S., 2021. Risk mitigation of corporate social performance in US class action lawsuits. Financial Analysts Journal 77 (2), 43–65.

Ferrell, A., Liang, H., Renneboog, L., 2016. Socially responsible firms. Journal of Financial Economics 122 (3), 585–606.

Ferris, S. P., Jayaraman, N., Liao, M.-Y. S., 2020. Better directors or distracted directors? An international analysis of busy boards. Global Finance Journal 44, 100437.

Fich, E. M., Shivdasani, A., 2006. Are busy boards effective monitors? The Journal of Finance 61 (2), 689–724.

Field, L., Lowry, M., Mkrtchyan, A., 2013. Are busy boards detrimental? Journal of Financial Economics 109 (1), 63–82.

Fiss, P. C., Zajac, E. J., 2006. The symbolic management of strategic change: Sensegiving via framing and decoupling. Academy of Management Journal 49 (6), 1173–1193.

Fombrun, C. J., Gardberg, N. A., Barnett, M. L., 2000. Opportunity platforms and safety nets: Corporate citizenship and reputational risk. Business and Society Review 105 (1), 85–106.

Frazzini, A., Israel, R., Moskowitz, T. J., 2018. Trading costs. SSRN Working Paper 3229719.

Frazzini, A., Pedersen, L. H., 2014. Betting against beta. Journal of Financial Economics 111 (1), 1–25.

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

Friedman, M., 1970. The social responsibility of business is to increase its profits. The New York Times Magazine, 13 September.

Fu, R., Tang, Y., Chen, G., 2020. Chief sustainability officers and corporate social (Ir)responsibility. Strategic Management Journal 41 (4), 656–680.

Gardberg, N. A., Fombrun, C. J., 2006. Corporate citizenship: Creating intangible assets across institutional environments. Academy of Management Review 31 (2), 329–346.

Gaur, S. S., Bathula, H., Singh, D., 2015. Ownership concentration, board characteristics and firm performance: A contingency framework. Management Decision 54 (5), 911–931.

Gentry, R. J., Shen, W., 2010. The relationship between accounting and market measures of firm financial performance: How strong is it? Journal of Managerial Issues 22 (4), 514–530.

Gillan, S. L., Starks, L. T., 2000. Corporate governance proposals and shareholder activism: The role of institutional investors. Journal of Financial Economics 57 (2), 275–305.

Godfrey, C., Hoepner, A. G., Lin, M.-T., Poon, S.-H., 2020. Women on boards and corporate social irresponsibility: Evidence from a Granger style reverse causality minimisation procedure. The European Journal of Finance, 1–27.

Godfrey, L. G., 1978. Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. Econometrica 46 (6), 1293–1301.

Godfrey, P. C., 2005. The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. Academy of Management Review 30 (4), 777–798.

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.

Gomanee, K., Girma, S., Morrissey, O., 2005. Aid, public spending and human welfare: Evidence from quantile regressions. Journal of International Development 17 (3), 299–309.

Goss, A., Roberts, G. S., 2011. The impact of corporate social responsibility on the cost of bank loans. Journal of Banking and Finance 35 (7), 1794–1810.

Grappi, S., Romani, S., Bagozzi, R. P., 2013. Consumer response to corporate irresponsible behavior: Moral emotions and virtues. Journal of Business Research 66 (10), 1814–1821.

Greening, D. W., Turban, D. B., 2000. Corporate social performance as a competitive advantage in attracting a quality workforce. Business & Society 39 (3), 254–280.

Grossman, S., 1976. On the efficiency of competitive stock markets where trades have diverse information. The Journal of Finance 31 (2), 573–585.

Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70 (3), 393–408.

GSIA, 2018. Global sustainable investment review. Global Sustainable Investment Alliance, Sydney.

Hackston, D., Milne, M. J., 1996. Some determinants of social and environmental disclosures in New Zealand companies. Accounting, Auditing & Accountability Journal 9, 77–108.

Hafsi, T., Turgut, G., 2013. Boardroom diversity and its effect on social performance: Conceptualization and empirical evidence. Journal of Business Ethics 112 (3), 463–479.

Hahn, T., Figge, F., Pinkse, J., Preuss, L., 2010. Trade-offs in corporate sustainability: You can't have your cake and eat it. Business Strategy and the Environment 19 (4), 217–229.

Halbritter, G., Dorfleitner, G., 2015. The wages of social responsibility–where are they? A critical review of ESG investing. Review of Financial Economics 26, 25–35.

Handelman, J. M., Arnold, S. J., 1999. The role of marketing actions with a social dimension: Appeals to the institutional environment. Journal of Marketing 63 (3), 33–48.

Harris, M., Raviv, A., 1990. Capital structure and the informational role of debt. The Journal of Finance 45 (2), 321–349.

Harris, M., Raviv, A., 2008. A theory of board control and size. The Review of Financial Studies 21 (4), 1797–1832.

Heal, G., 2005. Corporate social responsibility: An economic and financial framework. The Geneva papers on risk and insurance 30 (3), 387–409.

Heinkel, R., Kraus, A., Zechner, J., 2001. The effect of green investment on corporate behavior. Journal of Financial and Quantitative Analysis 36 (4), 431–449.

Hendricks, D., Patel, J., Zeckhauser, R., 1993. Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988. The Journal of Finance 48 (1), 93–130.

Hillman, A. J., Keim, G. D., 2001. Shareholder value, stakeholder management, and social issues: What's the bottom line? Strategic Management Journal 22 (2), 125–139.

Ho, F. N., Wang, H.-M. D., Vitell, S. J., 2012. A global analysis of corporate social performance: The effects of cultural and geographic environments. Journal of Business Ethics 107 (4), 423–433.

Hofstede, G., 2001. Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations. SAGE Publications, Thousand Oaks, CA.

Hofstede, G., Hofstede, G. J., Minkov, M., 2010. Cultures and organizations: Software of the mind. 3rd Edition. McGraw-Hill, New York.

Hofstede, G., McCrae, R. R., 2004. Personality and culture revisited: Linking traits and dimensions of culture. Cross-Cultural Research 38 (1), 52–88.

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

Huang, L.-Y., Lai, G. C., McNamara, M., Wang, J., 2011. Corporate governance and efficiency: Evidence from US property-liability insurance industry. Journal of Risk and Insurance 78 (3), 519–550.

Hull, C. E., Rothenberg, S., 2008. Firm performance: The interactions of corporate social performance with innovation and industry differentiation. Strategic Management Journal 29 (7), 781–789.

Ioannou, I., Serafeim, G., 2012. What drives corporate social performance? The role of nation-level institutions. Journal of International Business Studies 43 (9), 834–864.

Jain, T., Zaman, R., 2020. When boards matter: The case of corporate social irresponsibility. British Journal of Management 31 (2), 365–386.

Jamali, D., Mirshak, R., 2007. Corporate social responsibility (CSR): Theory and practice in a developing country context. Journal of Business Ethics 72 (3), 243–262.

Jensen, M. C., 1986. Agency costs of free cash flow, corporate finance, and takeovers. The American Economic Review 76 (2), 323–329.

Jensen, M. C., 1993. The modern industrial revolution, exit, and the failure of internal control systems. The Journal of Finance 48 (3), 831–880.

Jizi, M. I., Salama, A., Dixon, R., Stratling, R., 2014. Corporate governance and corporate social responsibility disclosure: Evidence from the US banking sector. Journal of Business Ethics 125 (4), 601–615.

Kagan, R. A., Gunningham, N., Thornton, D., 2003. Explaining corporate environmental performance: How does regulation matter? Law & Society Review 37 (1), 51–90.

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.

Karpoff, J. M., Lott, Jr, J. R., Wehrly, E. W., 2005. The reputational penalties for environmental violations: Empirical evidence. The Journal of Law and Economics 48 (2), 653–675.

Karpoff, J. M., Malatesta, P. H., Walkling, R. A., 1996. Corporate governance and shareholder initiatives: Empirical evidence. Journal of Financial Economics 42 (3), 365–395.

Kaufmann, D., Kraay, A., Mastruzzi, M., 2010. The worldwide governance indicators: Methodology and analytical issues. Policy Research Working Paper No. 5430, World Bank, Washington, DC.

Kempf, A., Osthoff, P., 2007. The effect of socially responsible investing on portfolio performance. European Financial Management 13 (5), 908–922.

Knox, S., Maklan, S., 2004. Corporate social responsibility: Moving beyond investment towards measuring outcomes. European Management Journal 22 (5), 508–516.

Knox, S., Maklan, S., French, P., 2005. Corporate social responsibility: Exploring stakeholder relationships and programme reporting across leading FTSE companies. Journal of Business Ethics 61 (1), 7–28.

Kotchen, M., Moon, J. J., 2012. Corporate social responsibility for irresponsibility. The B.E. Journal of Economic Analysis & Policy 12 (1), 1–21.

Kreander, N., Gray, R. H., Power, D. M., Sinclair, C. D., 2005. Evaluating the performance of ethical and non-ethical funds: A matched pair analysis. Journal of Business Finance & Accounting 32 (7-8), 1465–1493.

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

Lenz, I., Wetzel, H. A., Hammerschmidt, M., 2017. Can doing good lead to doing poorly? Firm value implications of CSR in the face of CSI. Journal of the Academy of Marketing Science 45 (5), 677–697.

Levinson, D. J., 2004. Empire-building government in constitutional law. Harv. L. Rev. 118 (3), 915–972.

Liang, H., Renneboog, L., 2017. On the foundations of corporate social responsibility. The Journal of Finance 72 (2), 853–910.

Lin-Hi, N., Müller, K., 2013. The CSR bottom line: Preventing corporate social irresponsibility. Journal of Business Research 66 (10), 1928–1936.

Lipton, M., Lorsch, J. W., 1992. A modest proposal for improved corporate governance. The Business Lawyer 48 (1), 59–77.

Lundgren, T., Olsson, R., 2009. How bad is bad news? Assessing the effects of environmental incidents on firm value. American Journal of Finance and Accounting 1 (4), 376–392.

Mallin, C. A., Michelon, G., 2011. Board reputation attributes and corporate social performance: An empirical investigation of the US best corporate citizens. Accounting and Business Research 41 (2), 119–144.

Manita, R., Bruna, M. G., Dang, R., Houanti, L., 2018. Board gender diversity and ESG disclosure: Evidence from the USA. Journal of Applied Accounting Research 19, 206–224.

Martínez-Ferrero, J., Banerjee, S., García-Sánchez, I. M., 2016. Corporate social responsibility as a strategic shield against costs of earnings management practices. Journal of Business Ethics 133 (2), 305–324.

Matten, D., Moon, J., 2008. "Implicit" and "explicit" CSR: A conceptual framework for a comparative understanding of corporate social responsibility. Academy of Management Review 33 (2), 404–424.

Mayer, R. C., Schoorman, F. D., 1992. Predicting participation and production outcomes through a two-dimensional model of organizational commitment. Academy of Management Journal 35 (3), 671–684.

McCabe, A. C., Ingram, R., Dato-On, M. C., 2006. The business of ethics and gender. Journal of Business Ethics 64 (2), 101–116.

McGuire, J. B., Schneeweis, T., Branch, B., 1990. Perceptions of firm quality: A cause or result of firm performance. Journal of Management 16 (1), 167–180.

McGuire, J. B., Sundgren, A., Schneeweis, T., 1988. Corporate social responsibility and firm financial performance. Academy of Management Journal 31 (4), 854–872.

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

McWilliams, A., Siegel, D., 2001. Corporate social responsibility: A theory of the firm perspective. Academy of Management Review 26 (1), 117–127.

Merton, R., 1987. A simple model of capital market equilibrium with incomplete information. The Journal of Finance 42 (3), 483–510.

Mervelskemper, L., Streit, D., 2017. Enhancing market valuation of ESG performance: Is integrated reporting keeping its promise? Business Strategy and the Environment 26 (4), 536–549.

Meyer, J. P., Allen, N. J., 1991. A three-component conceptualization of organizational commitment. Human Resource Management Review 1 (1), 61–89.

Meyer, J. W., Rowan, B., 1977. Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology 83 (2), 340–363.

Meznar, M. B., Nigh, D., 1995. Buffer or bridge? Environmental and organizational determinants of public affairs activities in American firms. Academy of Management Journal 38 (4), 975–996.

Minton, B. A., Taillard, J. P., Williamson, R., 2014. Financial expertise of the board, risk taking, and performance: Evidence from bank holding companies. Journal of Financial and Quantitative Analysis 49 (2), 351–380.

Mueller, K., Hattrup, K., Spiess, S.-O., Lin-Hi, N., 2012. The effects of corporate social responsibility on employees' affective commitment: A cross-cultural investigation. Journal of Applied Psychology 97 (6), 1186–1200.

Muhammad, A., Seale, J. L., Meade, B., Regmi, A., 2011. International evidence on food consumption patterns: An update using 2005 international comparison program data. USDA-ERS Technical Bulletin 1929.

Muñoz, F., Vargas, M., Marco, I., 2013. Environmental mutual funds: Financial performance and managerial abilities. Journal of Business Ethics 124 (4), 551–569.

Murphy, P. E., Schlegelmilch, B. B., 2013. Corporate social responsibility and corporate social irresponsibility: Introduction to a special topic section. Journal of Business Research 66 (10), 1807–1813.

Mynhardt, R., Makarenko, I., Plastun, O., 2017. Market efficiency of traditional stock market indices and social responsible indices: The role of sustainability reporting. Investment Management and Financial Innovations 14 (2), 94–106.

Narayanan, M. P., 1985. Managerial incentives for short-term results. The Journal of Finance 40 (5), 1469–1484.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.

North, D. C., 1990. Institutions, institutional change and economic performance. Cambridge University Press, New York.

Ohlson, J. A., 1995. Earnings, book values, and dividends in equity valuation. Contemporary Accounting Research 11 (2), 661–687.

Oliver, C., 1991. Strategic responses to institutional processes. Academy of Management Review 16 (1), 145–179.

Orbach, R., 2017. Bringing talent off the bench and into the game: The underrepresentation of women in the boardroom. Fordham Journal of Corporate & Financial Law 22 (2), 203–256.

Orlitzky, M., Schmidt, F. L., Rynes, S. L., 2003. Corporate social and financial performance: A meta-analysis. Organization Studies 24 (3), 403–441.

Ormiston, M. E., Wong, E. M., 2013. License to ill: The effects of corporate social responsibility and CEO moral identity on corporate social irresponsibility. Personnel Psychology 66 (4), 861–893.

Park, S.-B., 2018. Multinationals and sustainable development: Does internationalization develop corporate sustainability of emerging market multinationals? Business Strategy and the Environment 27 (8), 1514–1524.

Peloza, J., 2006. Using corporate social responsibility as insurance for financial performance. California Management Review 48 (2), 52–72.

Peterson, D. K., 2004. The relationship between perceptions of corporate citizenship and organizational commitment. Business and Society 43 (3), 296–319.

Pfeffer, J., Salancik, G. R., 1978. The external control of organizations: A resource dependence perspective. Harper & Row, New York.

Pintekova, A., Kukacka, J., 2019. Corporate social responsibility and stock prices after the financial crisis: The role of primary strategic CSR activities. SSRN Working Paper 3380881.

Porter, M. E., Kramer, M. R., 2002. The competitive advantage of corporate philanthropy. Harvard Business Review 80 (12), 56–69.

Post, C., Rahman, N., Rubow, E., 2011. Green governance: Boards of directors' composition and environmental corporate social responsibility. Business & Society 50 (1), 189–223.

Preston, L. E., O'bannon, D. P., 1997. The corporate social-financial performance relationship: A typology and analysis. Business & Society 36 (4), 419–429.

Price, J. M., Sun, W., 2017. Doing good and doing bad: The impact of corporate social responsibility and irresponsibility on firm performance. Journal of Business Research 80, 82–97.

Prior, D., Surroca, J., Tribó, J. A., 2008. Are socially responsible managers really ethical? Exploring the relationship between earnings management and corporate social responsibility. Corporate Governance: An International Review 16 (3), 160–177.

Qiu, Y., Shaukat, A., Tharyan, R., 2016. Environmental and social disclosures: Link with corporate financial performance. The British Accounting Review 48 (1), 102–116.

Rallapalli, K. C., Vitell, S. J., Wiebe, F. A., Barnes, J. H., 1994. Consumer ethical beliefs and personality traits: An exploratory analysis. Journal of Business Ethics 13 (7), 487–495.

Rao, K., Tilt, C., 2016. Board composition and corporate social responsibility: The role of diversity, gender, strategy and decision making. Journal of Business Ethics 138 (2), 327–347.

Refinitiv, 2021. Environmental, social and governance (ESG) scores from Refinitiv. https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf.

Renneboog, L., Ter Horst, J., Zhang, C., 2008. Socially responsible investments: Institutional aspects, performance, and investor behavior. Journal of Banking & Finance 32 (9), 1723–1742.

Revelli, C., Viviani, J.-L., 2015. Financial performance of socially responsible investing (SRI): What have we learned? A meta-analysis. Business Ethics: A European Review 24 (2), 158–185.

Reverte, C., 2009. Determinants of corporate social responsibility disclosure ratings by Spanish listed firms. Journal of Business Ethics 88 (2), 351–366.

Ringov, D., Zollo, M., 2007. The impact of national culture on corporate social performance. Corporate Governance 7 (4), 476–485.

Robinson, M., Kleffner, A., Bertels, S., 2011. Signaling sustainability leadership: Empirical evidence of the value of DJSI membership. Journal of Business Ethics 101 (3), 493–505.

Rodriguez, P., Uhlenbruck, K., Eden, L., 2005. Government corruption and the entry strategies of multinationals. Academy of Management Review 30 (2), 383–396.

Rodriguez-Dominguez, L., Gallego-Alvarez, I., Garcia-Sanchez, I. M., 2009. Corporate governance and codes of ethics. Journal of Business Ethics 90 (2), 187–202.

Rowley, T., Berman, S., 2000. A brand new brand of corporate social performance. Business & Society 39 (4), 397–418.

Rynes, S. L., Bretz Jr, R. D., Gerhart, B., 1991. The importance of recruitment in job choice: A different way of looking. Personnel Psychology 44 (3), 487–521.

Salancik, G. R., Pfeffer, J., 1978. A social information processing approach to job attitudes and task design. Administrative Science Quarterly 23 (2), 224–253.

Sánchez, C. M., 2000. Motives for corporate philanthropy in El Salvador: Altruism and political legitimacy. Journal of Business Ethics 27 (4), 363–375.

Schaltegger, S., Hörisch, J., 2017. In search of the dominant rationale in sustainability management: Legitimacy-or profit-seeking? Journal of Business Ethics 145 (2), 259–276.

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.

Schreck, P., Raithel, S., 2018. Corporate social performance, firm size, and organizational visibility: Distinct and joint effects on voluntary sustainability reporting. Business and Society 57 (4), 742–778.

Schunck, R., 2013. Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. The Stata Journal 13 (1), 65–76.

Scott, W. R., 1987. The adolescence of institutional theory. Administrative Science Quarterly 32 (4), 493–511.

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.

Servaes, H., Tamayo, A., 2014. How do industry peers respond to control threats? Management Science 60 (2), 380–399.

Shackleton, M., Yan, J., Yao, Y., 2022. What drives a firm's ES performance? Evidence from stock returns. Journal of Banking and Finance 136, 106304.

Sharfman, M. P., Fernando, C. S., 2008. Environmental risk management and the cost of capital. Strategic Management Journal 29 (6), 569–592.

Shleifer, A., 2004. Does competition destroy ethical behavior? American Economic Review 94 (2), 414–418.

Statman, M., 2000. Socially responsible mutual funds (corrected). Financial Analysts Journal 56 (3), 30–39.

Statman, M., Glushkov, D., 2009. The wages of social responsibility. Financial Analysts Journal 65 (4), 33–46.

Stein, J. C., 1988. Takeover threats and managerial myopia. Journal of Political Economy 96 (1), 61–80.

Strickland, D., Wiles, K. W., Zenner, M., 1996. A requiem for the USA: Is small shareholder monitoring effective? Journal of Financial Economics 40 (2), 319–338.

Suchman, M. C., 1995. Managing legitimacy: Strategic and institutional approaches. Academy of Management Review 20 (3), 571–610.

Tang, Z., Hull, C. E., Rothenberg, S., 2012. How corporate social responsibility engagement strategy moderates the CSR–financial performance relationship. Journal of Management Studies 49 (7), 1274–1303.

Thompson, J. B., 2005. The new visibility. Theory, Culture & Society 22 (6), 31–51.

Thomson Reuters, 2019. Environmental, social and governance (ESG) scores from Refinitiv. https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica 26 (1), 24–36.

Turban, D. B., Greening, D. W., 1997. Corporate social performance and organizational attractiveness to prospective employees. Academy of Management Journal 40 (3), 658–672.

Tyler, T. R., Lind, E. A., Huo, Y. J., 2000. Cultural values and authority relations: The psychology of conflict resolution across cultures. Psychology, Public Policy, and Law 6 (4), 1138–1163.

Ullmann, A. A., 1985. Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of US firms. Academy of Management Review 10 (3), 540–557.

US SIF, 2018. Report on US sustainable and impact investing trends 2018. Forum for Sustainable and Responsible Investment, Washington, DC.

US SIF, 2020. Report on US sustainable and impact investing trends 2020. Forum for Sustainable and Responsible Investment, Washington, DC.

Utz, S., Wimmer, M., 2014. Are they any good at all? A financial and ethical analysis of socially responsible mutual funds. Journal of Asset Management 15 (1), 72–82.

Valentine, S. R., Rittenburg, T. L., 2007. The ethical decision making of men and women executives in international business situations. Journal of Business Ethics 71 (2), 125–134.

Van der Laan, G., Van Ees, H., Van Witteloostuijn, A., 2008. Corporate social and financial performance: An extended stakeholder theory, and empirical test with accounting measures. Journal of Business Ethics 79 (3), 299–310.

Vasilescu, C., Wisniewski, T. P., 2019. Rebuilding a shattered reputation: Financial reporting quality and donations following corporate controversies. SSRN Working Paper 3484818.

Velasquez, M. G., Velazquez, M., 2002. Business ethics: Concepts and cases. Vol. 111. Prentice Hall, Upper Saddle River, New York.

Vitell, S. J., Nwachukwu, S. L., Barnes, J. H., 1993. The effects of culture on ethical decision-making: An application of Hofstede's typology. Journal of Business Ethics 12 (10), 753–760.

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

Waldman, D. A., De Luque, M. S., Washburn, N., House, R. J., Adetoun, B., Barrasa, A., Bobina, M., Bodur, M., Chen, Y.-J., Debbarma, S., et al., 2006a. Cultural and leadership predictors of corporate social responsibility values of top management: A globe study of 15 countries. Journal of International Business Studies 37 (6), 823–837.

Waldman, D. A., Siegel, D. S., Javidan, M., 2006b. Components of CEO transformational leadership and corporate social responsibility. Journal of Management Studies 43 (8), 1703–1725.

Wang, C.-J., 2012. Board size and firm risk-taking. Review of Quantitative Finance and Accounting 38 (4), 519–542.

Watts, R. L., Zimmerman, J. L., 1986. Positive accounting theory. Prentice-Hall, Englewood Cliffs.

Webb, E., 2004. An examination of socially responsible firms' board structure. Journal of Management and Governance 8 (3), 255–277.

Weick, K. E., Sutcliffe, K. M., Obstfeld, D., 2005. Organizing and the process of sensemaking. Organization Science 16 (4), 409–421.

Whitley, R., 1999. Divergent capitalisms: The social structuring and change of business systems. Oxford University Press, Oxford.

Williams, G., Zinkin, J., 2008. The effect of culture on consumers' willingness to punish irresponsible corporate behaviour: Applying Hofstede's typology to the punishment aspect of corporate social responsibility. Business Ethics: A European Review 17 (2), 210–226.

Wimmer, M., 2013. ESG-persistence in socially responsible mutual funds. Journal of Management and Sustainability 3 (1), 9–15.

Ye, K., Zhang, R., 2011. Do lenders value corporate social responsibility? Evidence from China. Journal of Business Ethics 104 (2), 197–207.

Yermack, D., 1996. Higher market valuation of companies with a small board of directors. Journal of Financial Economics 40 (2), 185–211.

Zubeltzu-Jaka, E., Álvarez-Etxeberria, I., Ortas, E., 2020. The effect of the size of the board of directors on corporate social performance: A meta-analytic approach. Corporate Social Responsibility and Environmental Management 27 (3), 1361–1374.