Impact of Environmental Performance on Firm Value: Evidence from Large US and European Union Firms

Master Thesis Department Finance

J.J.Sengers (SNR 1258865)

Supervisor: Professor F.S. Hosseini Tash

Second reader: Professor Dr. J.A. Crego

Date: 15th of August

TILBURG UNIVERSITY

Abstract

This study investigates the impact of environmental performance on stock prices in the short- and long-term by using the event study methodology to Newsweek's 'Green Ranking' announcements of 2014 until 2017 for large U.S. and European Union firms. To address the concern of cross-sectional dependence on the announcement date, a benchmark of 79 country-specific sector indexes are used. The results provide limited evidence, which indicates that high environmental performance is negatively related to the stock market performance in the short-term. Also, the findings show similar outcomes for subsamples for firms headquartered in the United States and the European Union. Moreover, this study finds no significant difference in the stock price reaction between the two geographical locations. After examining the short-term stock market reaction, the results over the nine-month holding period show that environmental performance is positively related to stock market performance. Especially the difference between the top- and bottom quintile of the green ranking shows significantly positive results in the nine-month holding period. This study indicates that there is a market incentive for companies from the United States and the European Union to improve their environmental performance in the long-term.

TILBURG UNIVERSITY

Table of contents

Preface

Dear reader,

Here it is. I present to you the final work, which hopefully completes my master Finance. By finishing this thesis, the last part of my time as a student has come. The past couple of months were intense but made me appreciate and understand academic research. First of all, I would like to thank Fatemeh Hosseini Tash for her help and interesting discussions during the process of this thesis. Looking back at the past couple of years as a student, it was an unforgettable journey where I took a lot of opportunities, made a lot of new friends and met my girlfriend Daphne. I want to thank every single one of them, and also my family, for always supporting me, motivating me and by making great memories during my student life! But here it is, my final academic accomplishment. Hopefully, it is an interesting reading experience.

Joost Sengers

Tilburg, July, 2019

I. Introduction

Corporate Social Responsibility (CSR) received growing attention over the past decades. Milton Friedman started the first real debate about CSR in 1962 when he argued forcefully that the belief of social responsibility is "fundamentally subversive" (Carroll, 1979). According to Flammer (2013), the initial focus of CSR was on the "social" responsibility (e.g., community-based programs, fair wages). However, the inclusion of environmental responsibility is a recent development (e.g., reduction of CO_2 emission). Nowadays, there are increasing concerns about the environment and the rising costs associated with environmental protection, which makes sustainable development a key concern facing the business world. (Meric, Watson and Meric, 2012). Moreover, *The Economist*[1] conducted an online survey in 2005, which stated that 85% of investors and executives believed CSR initiatives were significant determinants of investment decisions compared to only 44% in 2000. A more recent survey of 766 CEOs conducted by the United Nations Global Compact (UNGC) and Accenture found that 93 percent of the CEOs believed that sustainability is critical to future success (UNGC & Accenture, 2010). Also, the studies of Judge and Douglas (1998) and the New York Times (2007) stated that successful and competitive firms began to value their environmental performance. To examine the effect of the growing awareness of environmental performance, this research further investigates stockholders' actions toward corporate social responsibility and environmentalism.

This study is motivated by the theoretical model of Klassen and McLaughlin (1996), which predicts a positive relationship between improved environmental performance and improved financial performance due to avoiding penalties, improved productivity, secure competitive positioning, and higher product contribution margins. More recently, Yadav, Han, and Rho (2016) investigated this relationship by using the environmental performance indicators of Newsweek to investigate the impact of the green ranking announcement on the firm value of the companies listed on the ranking. They found a positive market reaction when Newsweek announces their green ranking, and they state that the market reaction is positively related to a companies green score. However, the findings of Yadav et al. (2016) are based on a single event and only by investigating United States companies in the short-term. To overcome these limitations, this study extends the

[1] The Economist is an English weekly newspaper which began publication in September 1843 and owned by the Economist Group in London. They offer authoritative opinion and insights on politics, international news, business, science, finance, and technology.

empirical model of method of Yadav et al. (2016) by using multiple years, both companies from the United States and the European Union, and looking at the short- and long-term impact of the Newsweek green score. Therefore, this study tries to find evidence for the following research question: *"What is the impact of Newsweek's Green Ranking announcement on firm value in The United States and The European Union?"*

This study uses the short- and long-term event study methodology to investigate the market reaction on the announcement of the Newsweek green ranking. The findings of this study are not in line with the findings of Yadav et al. (2016) since the short-term results show evidence of a market reaction, which is negatively related to a firms' green score. Moreover, this study does not find any difference in the reaction on the stock market between firms headquartered in the European Union and the United States in the short-term. The results are not robust across different tests for standardization and cross-sectional correlation, and therefore, the results have to be interpreted with caution. This study also executed the long-term event study approach, which measures the nine-month holding period of the stocks in the green scores for the total sample. The long-term findings show that good environmental performance leads to improved stock market performance. Additionally, the results support the theoretical framework of Klassen et al. (1996) as it finds that the 100 top-ranked firms perform significantly better than the 100 lowest-ranked firms. The results are robust by randomizing the ranking.

This paper uses an international sample of stocks included in the Newsweek's green score list, being companies from the European Union and the United States. The sample consists of daily data ranging from 2014 until 2017. Since 2014, the methodology of Newsweek remains unchanged, which allows to compare results of the rankings over these years. The sample contains large US and European Union firms since the Newsweek green score list contains the 500 largest companies based on revenue and market capitalization. This study uses Datastream, Compustat Global, Compustat North America, The Wall Street Journal, and Factiva as data sources.

The findings of this paper contribute to the growing empirical literature on the link between corporate environmental performance and corporate financial performance by showing the short-term reaction on the stock market when environmental performances are published in an international context. Moreover, consistent with the literature of Klassen and McLaughlin (1996), this research provides further evidence by showing that poor environmental performance leads to lower stock market performance in the long-term and that, on average, better environmental performance leads to improved performance on the stock market.

The remainder of this paper proceeds as follows. Section II discusses existing literature on the relationship between financial performance and environmental performance in an international context. Second, section III describes the data collection and summary statistics. Third, section IV outlines the empirical model used in this analysis and discusses robustness checks. Fourth, section V presents the results of the analysis. Finally, section VI states the conclusion, limitations, and recommends future research.

II. Literature Review and Hypothesis Development

This chapter provides the current state of the literature on the relationship between corporate environmental performance and financial performance as well as the geographical differences in the perception of environmental responsibility for companies. First, section 2.1 discusses the relationship between the overall corporate environmental performance and financial performance, and more specifically, the relationship with the overall firm value. Second, section 2.2 illustrates the differences in environmental disclosure and the effect on earnings forecasting. Furthermore, also the different perception of overall CSR is explained for firms from the United States and Europe. Finally, section 2.3 argues the mean-reverting ness of abnormal returns from an event study in the long term.

2.1 Corporate Environmental Performance on Firm Value

Corporate environmental performance (CEP) has no off-the-shelf definition, as Ilinitch et al. (1998, p.286) points out, “defining corporate environmental performance is not a straightforward task.” Trump et al. (2015) addressed this issue and found a clear and parsimonious definition of corporate environmental performance by looking at 16 articles with a precise definition of Corporate Environmental Performance. Finally, they concluded the term as “the results of an organization’s management of its environmental aspects” (ISO 1999).

This definition gives space to a broad understanding of CEP. Therefore, a lot of companies and industries use different environmental strategies to improve their CEP such as pollution control, the reduction of CO_2, reduction of energy consumption, recycling of products, and water reduction. The prevailing view of using those strategies and improving the CEP is that there is a fixed and inherent trade-off: ecology versus the economy (Porter et al., 1995). On the one side, the social benefits that arise from environmental standards and on the other side, the private costs for preventing pollution and other factors that could harm the environment. Porter et al. (1995) argue this static view by proving that adequately designed environmental standards can cause innovations that improve its value or lower the total cost of a product. Klassen and Mclaughlin (1996) extended Porters’ view where environmental performance leads to market gains as well as cost savings (Figure 1). Researches used these theories and the wide range of environmental strategies to investigate the impact on a firm’s financial performance as firms expect to benefit by differentiating products, increasing efficiency, building global compliances and reducing potential liabilities.

The model of Klassen and Mclaughlin (1996) states that excellent environmental performance affects the financial performance of the firm through both the revenue and cost side. On the revenue side, according to Rosewicz (1990), customers have a preference for environmental orientated firms. Ruf et al. (2001) even found that customers' are willing to pay a premium for the environmental attributes of products. More recent literature explains this willingness to pay a premium, as it gives satisfaction to the customer of having contributed to environmental improvement (Amato and Amato, 2012; Luo and Bhattacharya, 2006). Furthermore, the strong environmental performance offers a new basis for differentiation for the customer through environmental certifications. Finally, entering some markets in the long term makes it necessary for companies to improve their environmental performance due to environmental restrictions.

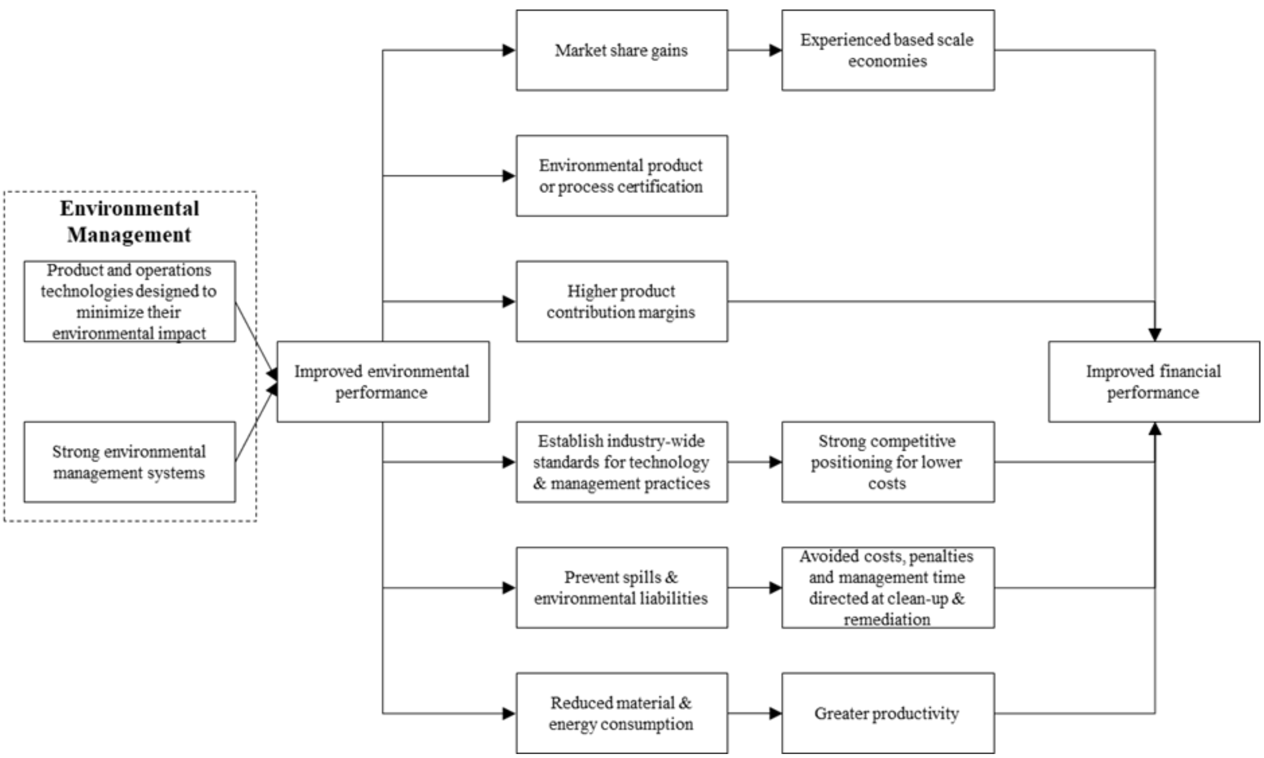


Figure 1. Model of Klassen and McLaughlin (1996) about the linkage of improved environmental performance to firm profitability.

In addition to potential revenue gains, improved environmental performance could also reduce cost. First of all, companies that invest heavily in environmental management systems through an improved production process can reduce energy consumption, intermediate material, and labor requirements, which reduces the overall production cost (Zeng et al., 2010). Secondly, Klassen and McLauglin (1996) state that companies can avoid the cost of future environmental crises, liabilities, and spoils due to the improvement in their environmental performance. Lastly, companies which move ahead of regulation to minimize the impact of their operations or products on the environment are in a better position to meet tighter standards in the future (Klassen and McLaughlin, 1996). Most of the time, current environmental requirements are based on the best available technology and therefore, Chynoweth and Kirschner (1993) state that companies that move ahead of regulation could gain a competitive advantage by establishing an industry standard and therefore create a potential barrier to entry.

The strand of literature that is most related to this study explains the relationship between CEP and a firms' stock price by using third-party evaluation. Yadav et al. (2016) recently published a paper about the impact of the 2012 Newsweek's Green Ranking[2] announcement on United States firms and found a positive effect on the performance in the stock market. However, using the 2009 Newsweek's green ranking, Meric et al. (2012) deduced no incentive for U.S. firms to go green and found a slightly negative impact on the six-month holding stock returns. On the other hand, Lyon and Shimshack (2015) found a positive impact of the environmental information disclosure on financial performance using a 2009 green ranking, where the top 100 firms, from the list of 500, showed a 0.6%-1.0% higher stock returns. Aditionally, Amato and Amato (2012) found a positive impact, using the same data as Lyon and Shimshack (2015), on the financial performance of firms in the top quartile. The overall conclusion of the impact of the Newsweek's Green Ranking announcement is rather positive, but there is still not a united answer by using the same source of data.

Other literature is in line with Porter et al. (1995) idea that environmental improvement can lead to positive financial performances. For instance, Russo and Fouts (1997) found a positive link between environmental performance and economic performance by analyzing 243 firms over two years, using independently developed environmental ratings. Furthermore, Klassen and McLaughlin (1996) find significant positive returns for firms which receive environmental

[2] Newsweek is an independent assessor which provides highly reliable information on the environmental performance of firms, covering a large number of investors through online and off-line media (Lyon and Shimshack, 2012). It consists about 10 differenct environmental variables that combine together an overall green score.

performance awards. These findings suggest that investors believe that environmental improvement will lead to market gains or cost reductions.

On the other hand, previous literature also finds a negative effect of improving environmental performance. The theory of Mathur and Marthur (2000) for example, they find a negative impact from the announcement of green marketing activities on stock value. Furthermore, Keele and DeHart (2001) and Cañón-de-Francia and Garcés-Ayerbe (2009) find a negative impact in respectively participating in EPA's[3] climate leaders' partnership and receiving ISO 1400I certification[4] on firms' market values. These results are in line with the prevailing view that improving the CEP is an inherent and fixed trade-off between social benefits and private costs. These findings will lead investors to believe that green marketing initiatives increase cost without sufficient offsetting increases in revenues.

To investigate the effect of corporate environmental performance on a firms' stock price through an event study, the theory of Malkiel and Fama (1970) about the efficient-market hypothesis has to be clear. Malkiel and Fama (1970) formulate the efficient-market hypothesis as 'a market is (informally) efficient if prices fully reflect all available information.' In this theory, three different market forms explain "all available information." At first, all available information only reflects the past prices and volume data; this is considered the weak form. Furthermore, the semi-strong form states that all available information consists of all public data for investors. Finally, if the market is in strong form, the prices reflect all information, including insider information. Recent literature of Cohen and Frazzini (2008) show that stock prices do not incorporate news involving related firms, which makes it possible to generate predictable consecutively price moves. This rejects the semi-strong form of all available information as the information was publicly available. The findings of Cohen and Frazzini (2008) leave space for abnormal returns as a reward for gathering and acting on costly information. Therefore, this research will act on the weak form of market efficiency, as stated by Malkiel and Fama (1970) and leaves space for moves in stock price by releasing information of third-party evaluation.

[3] The United States Environmental Protection Agency (EPA) is an independent agency of the federal government for environmental protection.
[4] ISO 14001 is a voluntary international standard that specifies requirements for an effective environmental management systems.

Overall, previous research finds positive and negative effects of corporate environmental performance on financial performance. Since these findings are not comparable to other studies because there are multiple approaches to determine corporate environmental performance, the results have to be interpreted with caution. This study uses the approach of Newsweek to determine environmental performance as it consists of eight key performance indicators and even ten indicators in 2017. Therefore, this paper will extend the model of Yadav et al. (2016). The list contains the largest global publicly-traded companies based on their market capitalization and revenue. Hence, the first hypothesis is formalized to test the positive relationship between CEP and financial performance since the information on environmental performance is meaningful to investors. This study uses the stock price in the research since Fama et al. (1969) argue that firms stock price is the best, unbiased estimate of firm value. Since firms are on the Newsweek's green score list based on their market capitalization and revenue, but not their environmental performance, only positive cumulative abnormal returns (CAR) for the top-ranked firms are expected. Amato et al. (2012) prove this as they state that top listed firms of the green score show stronger and more significant positive financial performance. Therefore, the following hypothesis is formalized to test the CARs of the best 100 companies on the Newsweek's green score list to prove that excellent environmental performance leads to positive stock price movement after the announcement.

Hypothesis 1: Top listed firms on the Newsweek green score show positive cumulative abnormal returns around the announcement of Newsweek's green score list.

According to the findings of Lyon et al. (2015), positive financial performance in the top 100 of the list experience abnormal returns, which are 0.6%-1.0% higher than returns of firms in the bottom 400. These results show that the green score of the firm on Newsweek's list is related to the cumulative abnormal returns. Therefore, the second hypothesis will test if the CARs are related to a firm's green score.

Hypothesis 2: CARs around the announcement date are positively related to a firm's green score.

2.2 International environmental disclosure and the effect of an earnings surprise

In addition to the actual environmental performance, some firms choose to disclose information regarding their environmental practices themselves, and this is called environmental disclosure.

According to the findings of Aerts et al. (2008), there is a positive relationship between enhanced environmental disclosure and more precise earnings forecasts by analysts. Furthermore, Brown et al. (1995) stated that analysts' earnings forecasts are relevant for their stock price forecast. Therefore, environmental disclosure is an important factor in forecasting a companies stock price.

According to Roe (2003), two significant determinants drive environmental disclosure for firms; these are both financial markets' and public interest considerations. On the one hand, firms disclose their environmental activities due to pressure from the financial markets. On the other hand, public pressure upon a firm forces management to disclose environmental information. Both parties, the stock market, and the media have significant control over a firm's future survival by either controlling its access to capital or its reputation and standing within the community (Aerts et al., 2008). Furthermore, Guthrie and Parker (1990) and Adamas and Kuasirikn (2000) found that environmental disclosure depends on the social consciousness of the country it comes from, taking into account the legal, sociopolitics, cultural and financial contexts. The country-specific public interest is, therefore, a significant driver of the amount of environmental disclosure a firm gives.

As stakeholders identities and interests vary cross-nationally, societies have developed different political, cultural, and financial system reflecting their ethics, social relations, and institutions (Whitley, 1999). Therefore, there are different environmental rules across continents. Saida (2009) states that both the United States, as well as the European Union (EU), have the most rigid environmental regulations in the world. The United State Congress, for example, developed a program to handle the ecological risks due to chemical gases. The EU also revealed rules to find out a future integration of environmental responsibility (Saida, 2009). Comparing both markets, Matten and Moon (2008) argue that in liberal market economies like the United States, disclosure of CSR activities is more common than in Scandinavian and Continental European countries. Bartolomeo et al. (2000) and Thomas et al. (1997) concluded that the level of environmental disclosure in the United States is larger than that of Australian and English companies. Also, the information on the environment is effectively related to the American judicial body, according to De La Bachellerie and Boillet-Mongodin (1993). These findings suggest that the environmental disclosure of American multinationals is superior to the European's. However, Saida (2009) questioned the results of Thomas et al. (1997) by finding that European multinationals reveal more environmental information than American multinationals. These findings suggest that, despite the more restricted laws in the United States, countries in the European Union disclose more ecological information voluntary.

Matten and Moon (2008) explain those voluntary environmental disclosures in the EU by expressing a conceptual framework for understanding differences in CSR across countries. They identify two different elements of CSR: the explicit- and the implicit CSR. Explicit CSR refers to corporate policies that articulate and assume responsibility for some societal interests. It consists of voluntary corporate policies, strategies, and programs which combine business and social value (Matten and Moon, 2008). The perceived expectations of different stakeholders of the corporation motivate the incentives and opportunities. On the other hand, implicit CSR describes corporations' role within the broader informal and formal institutions for society's concerns and interests. It consists of norms, values, and rules that result in requirements for corporations, and the role every major group in the society has to play. It requires to address stakeholder issues in collective rather than individual terms. Matten and Moon (2008) stated that the United States, as operating in a more liberal market, therefore, has explicit CSR in contrast to implicit CSR in Europe. Furthermore, explicit CSR disclosure is often the result of a strategic decision as it provides them with an opportunity to distinguish themselves. This is not the case for Continental European firms, as CSR is not a voluntary and intentional corporate decision in those countries, but more a reaction to a corporation's institutional environment (Matten and Moon, 2008).

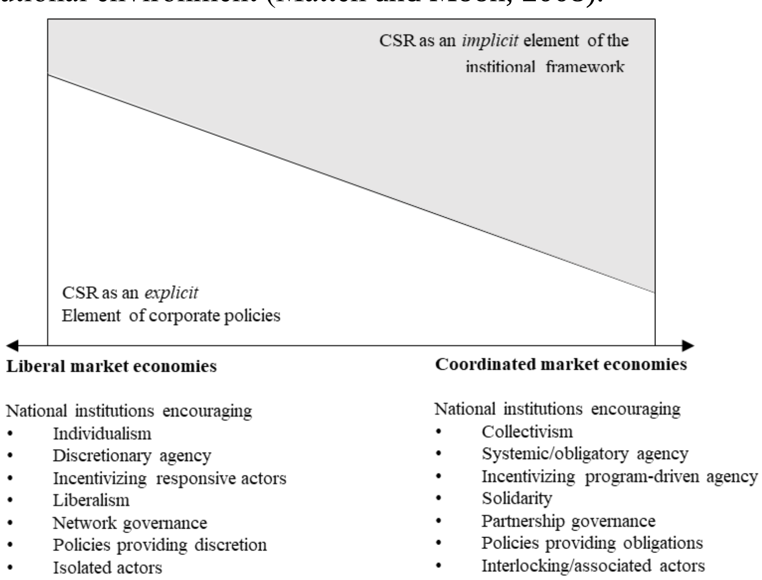


Figure 2. The model of Matten and Moon (2008) about two different elements of CSR: implicit and explicit CSR for different market economies.

To summarize, the literature provides empirical evidence about the difference in the amount of environmental disclosure between the United States and European countries. According to the research, the United States companies generally disclose more environmental information in comparison with the European Union. Furthermore, the framework of Matten and Moon (2008) gives clearer insides into the different environments of CSR expectations in both markets. In European countries, society expects companies to take their responsibility according to their environmental contribution in contrast to firms from the United States, where CSR activities are more often a strategic decision. According to Aerts et al. (2008), environmental disclosure leads to more precise earnings forecast, and this again leads to better stock price forecasts (Brown et al., 1995).

Since there are differences in the amount of environmental disclosure and the environmental perception across Europe and the United States, this study will look at the different reaction of the stock market when Newsweek announces their green scores. As the United States' environmental disclosure is higher compared to European environmental disclosure (Bartolomeo, 2000; Thomas et al., 1997) and the CSR in the United States is more explicit (Matten & Moon, 2008), this paper investigates if the stock price of the European companies will move more significantly by the announcement. Since the European Union has its own environmental legislation[5], only the companies headquartered in the European Union are part of the research. Taking both findings into account, the following hypothesis is formalized in the following context:

Hypothesis 3: CARs around the announcement date are more significantly related to a firm's green score for European Union firms.

2.3 under- and overreaction to post- and pre-event abnormal returns

According to the literature of De Bondt, and Thaler (1985), most people overreact to unexpected and dramatic news events. Their research found long-term return reversal from past winners and past losers and attributed this finding to the fact that investors give too little weight that performance tends to mean-revert. Furthermore, the paper of Fama (1998) state that there is developing literature that challenges the assumption that short returns windows have any lag in the response of prices, and the event would be short-lived. This literature suggests that stock prices

[5] The European Union's environmental legislation addresses issues such as sustainable energy, noise pollution, waste and water pollution, air quality, the thinning of the ozone layer and acid rain.

adjust slowly to information, and therefore, researches must extend the horizon of stock prices to get a full view of market inefficiency.

First of all, two behavioral theories support the theory of De Bondt & Thaler (1985) about the under- and overreaction to events. The first theory of Barberis, Shleifer, and Vishny (1998) come with evidence from two cognitive psychology judgment biases. First of all, when new evidence arises, the models explaining the results update too slowly, according to Edwards (1968). Second, Kahneman and Tversky (1982) argue that the recent patterns in the data get too much weight from investors, and they give too little weight to the properties of the population that generates the data. Overall, the theory of Barberis et al. (1998) states that there is a long-term return reversal. The second theory from Daniel, Hirshleifer, and Subramanyam (1997), state that the market consists of informed and uninformed investors. Here, only the informed investors have a bias through overconfidence and self-attribution. This makes them overreact to private information and underreact to public information in the short term. Eventually, this leads to long-term reversals as the public information overwhelms the bias.

Finally, Fama (1998) checked if the long-term under- and overreaction to information suggest market inefficiency. He concludes with a definite 'no' as underreaction is as common as an overreaction. This conclusion, however, does not mean that the investors do not include behavioral biases in their decisions as it states that the market efficiency theory still holds. As stated above, the theory of De Bondt and Thaler (1985) suggests that most people overreact to dramatic and unexpected news events. Therefore, this paper investigates if the event of the Newsweek's green score list leads to an over- or underreaction on the market. According to the behavioral theories of Barberis et al. (1998) and Daniel et al. (1997), this means that the long-term returns would be mean-reverting. For this research, the over- and underreaction would mean that the cumulative abnormal returns as a result of the publishment of the green score would decrease (increase) and go down (up) as time goes on. Taking all the theories into account, the final and last hypothesis of this research is as follows.

Hypothesis 4: The BHARs as a result of the Newsweek green score list are mean-reverting in the long-term.

III. Dataset and Institutional Details

This chapter provides an overview of the data used for this research. First, section 3.1 provides the sample selection procedure. Further, section 3.2 elaborates the main variables and explains the procedure on how to compute those variables. Third, section 3.3 discusses the control variables deemed relevant to the analysis used in this study. Finally, section 3.4 provides descriptive statistics on all the variables used in the empirical analysis of section 4 and 5.

3.1 Sample selection

The sample selection for the research in this paper comes from the American weekly news magazine *Newsweek* from the period 2014 until 2017. The magazine publishes a list of the corporate environmental performance of the 500 largest publicly-traded companies in the United States as well as the 500 largest publicly-traded companies globally both based on their revenue (2014 and 2015) and their market capitalization (2016 and 2017). *Newsweek* collaborated with Corporate Knights Capital[6] and HIP Investor[7] to complete the *Newsweek* Green Rankings. The companies scores are based on their performance on eight specific indicators what finally results in one overall green score. Since 2014, the methodology for calculating the overall green score is the same for seven of the eight key environmental performance indicators (KPI) over the sample period. The difference is the replacement of the reputation score in 2014 with the green revenue score in 2015 and 2016. Furthermore, since 2017, companies could get a deduction on their score by either a monetary fine or receive revenue through products or services that have a harmful impact. Section 3.2 explains the KPI's in more detail.

This research examines the effect of the announcement of the green scores from *Newsweek* on a firms' stock price in Europe and the United States. For Europe, the countries within the European Union are part of the sample of this research. Therefore, companies which have their headquarters' located in, for example, Switzerland or Russia are removed from the sample. For the United States, this paper only uses the companies listed on the global ranking to avoid factors influencing the results by only being part of the U.S. 500 ranking. From the possible total sample of 1,247 firms on the Newsweek Green Ranking list between 2014-2017 in the United States and

[6] Corporate Knights Capital is an independent research company based in Toronto, Canada. The parent company, Corporate Knights Inc., conducts corporate sustainability rankings since 2005.
[7] HIP Investor rates companies and investments for their benefit or cost to society, and the associated risks or opportunities for investors.

European Union, a total sample of 1,176 firms is collected at this point. The reason for neglecting the other 71 firms is due to not enough data in the databases accessible.

Furthermore, this paper uses The Wall Street Journal[8] and Factiva[9] to find other events three days before, and three days after the announcement of the Newsweek green score list of firms in the total sample, which could eventually affect the companies stock price in the event window. Therefore, another 43 companies are deleted from the sample, and this brings the total sample to 1,133 companies. Moreover, this paper excludes financial firms due to the different definition of variables for financial firms, as stated by Fama and French (1992). Therefore, an additional 228 firms are deleted from the sample, which brings the total sample to 905 firms over four years.

3.2 Main variables

This section discusses the essential variables used in the empirical analysis and how to compute them. First, the paper elaborates the data from *Newsweek* and the different indicators the data includes. Hereafter, section 3.2.2 elaborates the computation of the returns.

3.2.1 Newsweek's data

The essential variable in the empirical analysis is the overall green score. Amota and Amota (2012) site that the unique aspect of *Newsweek's* publication is the widespread dissemination of this information to the general public in an easily obtained and easily understood format and the data of *Newsweek* is also of high quality according to Lyon and Shimshack (2015). The methodology of the green score is transparent and rules-based and meets the test of being replicable by a third party. The overall score is a weighted average of their eight different key environmental performance indicators scores, which can add up to a total score of 100 percent. However, in 2017, there are two additional indicators which can distract point to the overall score. Furthermore, looking at the calculation of the key performance indicators (KPI), the first four KPI's have almost the same derivation. The methodology of the first KPI is explained more broadly since the second until the fourth KPI is calculated quite similar.

[8] The Wall Street Journal is United States' business-focused international daily newspaper located in New York City and published six days a week by Dow Jones & Company, which is a division of News Corp. It is one of the largest newspapers in the United States and has won 37 Pulitzer Prizes.

[9] Factiva is a research and business information tool owned by Dow Jones & Company, a division of News Corp. The tool provides search, dissemination, alerting, and other information management capabilities.

The first KPI used by the Newsweek's green score is the 'Combined Energy Productivity Score.' This indicator, which has an overall weight of 15% on the total score, is calculated in three steps. First, they calculate the energy productivity by dividing the total revenue with total energy consumption. This energy productivity is then percent-ranked against the energy productivity of all the peers in the industry group and multiplied by 0.75. In the second step, the change in the companies energy productivity from the past two years is calculated and also percent-ranked against the same industry group. The ranking is set into quartiles and multiplied by one till zero, according to the results, and then by 0.25. The final step adds step one and two, which finally gives the overall 'Combined Energy Productivity.'

The 'Combined Greenhouse Gas (GHG) Productivity' (weight 15%) is the second KPI and also defined with the same methodology as the Combined Energy Productivity. The GHG productivity is calculated as the total revenue divided by the total Greenhouse Gas Emissions (CO_2)[10] .

The third KPI is the 'Combined Water Productivity' (weight 15%) and is calculated by dividing the total revenue by the total water use per cubic meter (m^3) of the firm.

Furthermore, the following KPI is 'Combined Waste Productivity' (weight 15%). Companies can have a higher score when they recycle their waste as the score is calculated as the total revenue divided by the total waste generated (metric tonnes). The recycled waste of the company subtracts the total waste in the denominator. Further calculations of the fourth indicator are similar to the other three KPI's to get the overall indicator score.

The fifth indicator is the 'Green Revenue Percent Range' (weight 20%). This KPI has the most weight attached for the overall green score and is therefore considered most important according to *Newsweek*. The Green Revenue Score analysis the revenue of various segments of the company to determine the percentage of companies revenue that is green for each segment. The segments are weighted averages of the total revenue and add up to a total of 100%. For the 2017 Newsweek green score list, they used a percentage range to display the results of the Green

[10] Greenhouse Gas Emissions only includes scope 1 and scope 2 according to the GHG Protocal. GHG Protocal supplies the world's most widely used greenhouse gas accounting and reporting standards. Scope 1 and 2 are the direct emissions from owned or controlled sources and the indirect emissions from the generation of purchased energy.

Revenue in one of five ranges from 0-20% to 81-100%. Furthermore, Newsweek used 'Reputation' as a KPI in 2014 instead of 'Green Revenue.' This score is made up of two components: the RepRisk Index, which quantifies a companies exposure to environmental and captures the criticism, social, and governance risks. Secondly, a component is based on the number of environmental issues a company had during the year of investigation.

The other three indicators can either have a score of 0% or 100%. First, when a company has a sustainability pay link (weight 10%), which means that a senior executive team members' remuneration is linked with the achievement of environmental performance targets, the company gets a 100% score for this KPI. Moreover, the following indicator is the 'Sustainability Board Committee' (weight 5%), and a company receives a score of 100% when a committee at the Board of Directors his mandate is related to the sustainability of the firm. Finally, a company receives the full score with a weight of 5% when a third party does the audit of environmental metrics.

3.2.2 Daily Returns and Benchmark Returns

Another essential variable in this analysis is the calculation of the returns of each company on the Newsweek's green score list. The returns for the United States were conducted out of CRSP Daily Stock from the Wharton Research Data Services (WRDS), using the 'holding period return.' These holding period returns are then converted to log returns. The collecting period of the returns is from six months before the announcement until nine months after the announcement to check for the long term effect. The data from COMPUSTAT in WRDS does not include the 'holding period return' for European Union companies, and therefore, another measure for calculating the returns is used. Additionally, the database of COMPUSTAT does not include a closing price which takes into account stock splits and dividend payments. For this reason, the following formula is used, which takes into account both stock splits and dividend payments and is, therefore, a sufficient indicator to calculate returns.

$$Daily\ Adjusted\ Closing\ Price_{i,t} = \left(\frac{Daily\ Price_{i,t}}{Daily\ Adjustment\ Factor_{i,t}}\right) * DTRF_{i,t} \tag{3.1}$$

In this formula, the Daily Price is the closing price, not adjusted for stock splits and dividends, of a company. The Daily Adjustment Factor is the factor which converts the historical shares data to an equivalent current basis. Finally, the DTRF is the Daily Total Return Factor includes cash equivalent distributions along with reinvestment of dividends and the compounding influence of dividends paid on reinvested dividends.

Furthermore, to calculate the daily log returns, the Daily Adjusted Closing Price for European Union companies and the Holding Period Return for companies from the United States are divided by the Daily Adjusted Closing Price or Holding Period Return the day before. The following formula states the calculation of the returns.

$$Return_{i,t} = \log\left(\frac{Daily\ Adjusted\ Closing\ Price_{i,t}}{Daily\ Adjusted\ Closing\ Price_{i,t-1}}\right) \tag{3.2}$$

Finally, the country-specific sector index is used as a benchmark to compute the abnormal returns in the event study for the short-term period. Newsweek adds the Global Industry Classification Standard[11] (GICS) sectors for every firm listed on the Newsweek green score list, and therefore, these sectors are used for the benchmark return. The price index is collected from Datastream, and the log returns are calculated by dividing the log of the price index at time t with the log of the price index at time $t-1$. For the long-term event study, the five-factor model from Fama and French (1993) is used. For this analysis, the European and North American market returns are used in the formula. Section 4.3.1 further explains this in more detail.

3.3 Control Variables

Based on prior research concerning the relationship of corporate environmental performance and the overall firm value, multiple control variables are selected. These control variables control for other effects than the corporate environmental performance that affect the total firm value. All the control variables are taken from the end of the year before the Newsweek green score announcement and collected from the COMPUSTAT database. Capon, Farley, & Hoenig (1990) did a meta-analysis about the determinants of financial performance and found some factors that affect the financial performance of a firm which will be stated more broadly in this section.

[11] Standard & Poor and MSCI developed the Global Industry Classication Standardin 1999 as an industry taxonomy for the use of the global financial community. The GICS consists of 11 sectors, 24 industry groups, 69 industries and 158 sub-industries.

First of all, Yadav et al. (2016) found a positive and significant effect of the return on assets and the firm's stock price. For this reason, this paper will include return on assets as a control variable, and this is calculated as the ratio of net income to total assets.

Furthermore, growth in assets and sales is individually positive related to performance at both industry and firm-business levels of analysis, according to Capon et al. (1990) by analyzing 88 studies. This research uses the approach of Russo & Fouts (1997) for measuring a firm's growth rate as the annual change in sales, expressed as a percentage.

Thirdly, the overall size of the firm is another control variable used in this research. According to Russo and Fouts (1997), size is another causal variable which is prevalent in prior studios of performance. In this study, size is the natural logarithm of sales volume.

Fourtly, Opler et al. (1994) found that highly leveraged firms lose significant market share to their more conservatively financed competitors in industry downturns. They found that firms in the top leverage decile in their industries, see a decline in sales with an average amount of 26 in comparison with firms in the bottom leverage decile. For this reason, leverage is used as a control variable in the regression.

Finally, Russo and Fouts (1997) investigated the effect of industry growth on the positive relationship between environmental- and economic performance. They concluded that industry growth strengthens the relationship as it 'pays, even more, to be green' in high growth industries. Furthermore, Capon et al. (1990) also found that industry concentration is another determinant of firm performance. They investigated almost 100 studies, and over 1100 tests show a clear directional positive effect between industry concentration and firm performance. As the total sample consists of too many industries to get a correct overview, the sectors are used as a control variable and included as a dummy in the analysis of this paper. The Global Industry Classification Standard (GICS) is used for 'sectors.'

3.4 Descriptive statistics

This section provides the descriptive statistics of all the variables used in the paper. Tables A.1 present the number of observations, the percentage of the total sample, average green score, and the average cumulative abnormal returns for the distribution across countries, sectors, and years, respectively. Table A.1.1 in the Appendix reports the final dataset of companies for each country.

The number of companies from the United States is almost two-thirds of the total sample in comparison to companies from the European Union. Furthermore, European Union companies have, on average, higher green scores in comparison to companies from the United States with respectively 0.515 and 0.387. Moreover, the average cumulative abnormal return in the US is lower for firms in the sample. Table A.1.2 in the Appendix displays additional information regarding the sector-specific average green scores and average daily returns. The Energy Sector scores the lowest Average green score with 0.327. This is not surprising, as the Union of Concerned Scientists[12] states that producing electricity and power is the single largest source of global warming in the United States nowadays. On the other hand, Telecommunication Services is the 'greenest' sector, according to Table A.1.2, with an overall average score of 0.536. Finally, Table A.1.3 displays the distribution from the total sample for each year. The number of observations slowly falls over the sample period but stays at a minimum of 200 observations in 2017. The rising number of companies outside the United States and European Union into the global 500 lists of Newsweek causes the decrease, especially companies headquartered in Asia.

Table 1 presents the descriptive statistics of the dependent-, control-, dummy- and interaction variables of this paper. In Table 1, the number of observations, mean value, standard deviation, and several percentiles are displayed for the whole sample. The overall sample is well distributed as the mean and median are close to the central tendency of the Newsweek green score list. The average Newsweek green score of the sample is 0.43 and can be seen as higher than average, as the overall average score of the 2000 companies on the Newsweek's list has an average score of 0.39. Furthermore, one component of the green score has changed over the years and is displayed as 'Reputation' in 2014 and 'Green Revenue' from 2015 until 2017. As explained above, the Green Revenue score in 2017 is provided as a percentage range and is therefore transformed into a measure between 0.1 and 0.9. The sample has a positive cumulative abnormal return over the sample period of 0.09%.

[12] The Union of Scientists (UCS) is a United States nonprofit science advocacy organisation. It contains over 200,000 members including professional scientists.

Table 1
Sample Descriptive Statistics

This Table presents the descriptive statistics for the dependent-, control-, dummy- and interaction variables used in the overall study over the period 2014-2017. The number of observations, mean, standard deviation, minimum, 5th percentile, median, 95th percentile, and maximum are displayed respectively.

Variables	# Obs.	Mean	Std. Dev.	Min.	5th Perc.	Median	95th Perc.	Max.
Cumulative abnormal return	905	0.00	0.03	-0.14	-0.03	0.00	0.04	0.26
Rank	905	222	140	1	15	209	466	500
Newsweek Green Score	905	0.43	0.20	0.00	0.07	0.45	0.75	0.90
Energy Productivity	905	0.35	0.29	0.00	0.00	0.33	0.85	1.00
Carbon Productivity	905	0.41	0.27	0.00	0.00	0.40	0.85	1.00
Water Productivity	905	0.35	0.31	0.00	0.00	0.34	0.86	1.00
Waste Productivity	905	0.30	0.31	0.00	0.00	0.20	0.86	1.00
Reputation	236	0.38	0.28	0.00	0.00	0.33	0.90	1.00
Green Revenue ('15-'16)	469	0.56	0.23	0.02	0.08	0.60	0.89	0.95
Green Revenue ('17)	200	0.12	0.66	0.10	0.10	0.10	0.30	0.00
SustainabilityPay Link	905	0.67	0.47	0.00	0.00	1.00	1.00	1.00
Sustainability Themed Committee	905	0.68	0.47	0.00	0.00	1.00	1.00	1.00
Audit Score	905	0.67	0.47	0.00	0.00	1.00	1.00	1.00
ROA	888	0.06	0.06	-0.16	-0.01	0.06	0.16	0.31
Growth	885	0.03	0.16	-0.52	-0.17	0.02	0.26	0.88
Log(Size)	886	10.26	1.04	6.85	8.57	10.19	11.94	13.02
Leverage	888	0.93	2.55	-21.98	0.01	0.66	3.65	12.18
EU	905	0.35	0.48	0.00	0.00	0.00	1.00	1.00
EU * Green Score	905	0.18	0.27	0.00	0.00	0.00	0.72	0.90
Dummy_Deterioration	669	0.34	0.47	0.00	0.00	0.00	1.00	1.00
Dummy_New	669	0.22	0.41	0.00	0.00	0.00	1.00	1.00
Dummy_Improvement	669	0.44	0.50	0.00	0.00	0.00	1.00	1.00
Δ Green Score	528	0.01	0.12	-0.47	-0.20	0.01	0.21	0.40

The descriptive statistics of the control, dummy, and interaction variables are also displayed in Table 1. Due to missing data in COMPUSTAT Global and North-America, the total number of observations is smaller than the overall sample. To minimize the influence of outliers in our data, a 1% winsorization is used which transforms the top 0.5%, and the bottom 0.5% into the 99.5th percentile and the 0.5th percentile, respectively. The control variables have, on average, a positive return on assets and annual growth of 6% and 3% respectively. For both variables, the median is almost the same as the average, which indicates that the distribution is symmetric. Furthermore, the average logarithm of sales is 10.26, which is almost the same as the median of 10.19 which indicates that the sales is symmetrically distributed. Finally, the average leverage ratio is 0.93, which indicates that the total sample has on average less debt than equity. However, the median leverage ratio is 0.66, and this indicates that some outlines mainly drive the overall average. The extremely negative leverage ratio of -21.98 indicates that some firms have a negative stockholders' equity, which is small, and this results in extremely low leverage. The 95th percentile leverage ratio of 3.65 indicates that at least 5 percent of the companies in the sample is highly levered. Consistent with Tabel A.1.1, around 35% of the sample consist of European Union companies. Also, the interaction variable between Europe and the green score is displayed in Table 1 and do not show a disrupting result. Furthermore, the dummy variables for deterioration, new entrance, and improvement show that 44 percent of the subsample improved their green score, as 34 percent found a decrease in their green score. Finally, the difference in the green score for a company in successive years is displayed as Δ Green Score. On average, companies receive a 1% higher green score in the following year. However, this varies from a 47% downgrade to a 40% increase.

IV. Empirical Model

This chapter describes the methodology to test the hypotheses used in this paper. First, section 4.1 describes the event study used for calculating the abnormal returns from the Newsweek's green score list. Second, section 4.2 presents the model to evaluate whether the cumulative abnormal returns from the announcement of the green score list are significantly related to their overall green score. Third, section 4.3 describes the model used for testing the different perceptions of the announcement for European Union firms and United States' firms. Fourth, section 4.4 explains the model to test if the abnormal returns are mean-reverting in the long-term. Finally, section 4.5 provides additional robustness tests to verify the obtained results.

4.1 Event study and Cumulative Abnormal Returns

This study aims to investigate the links between market performance and environmental performance by estimating a market reaction to the announcement of the environmental performance given by Newsweek through an event study. This section further explains the event study and the calculations used to test the hypothesis formulated in chapter 2.

4.1.1 Definition of event study and identifying the event window

An event study measures the particular market returns associated with a specific event using financial market data. Since Fama et al. (1969) introduced the first event study, this methodology is widely used in research examining financial implications of events in various disciplines, such as marketing, strategic management, account and finance (Yadav et al., 2016). This methodology is useful to estimate the economic impact of the event by observing the reaction of the securities return on the stock market over a short period.

An example from the paper of Fama et al. (1969) as they address the methodology of an event study by comparing holding returns of stocks around an event date with the expected return if there were no event. The difference between the expected returns and the actual returns are defined as abnormal returns. For conducting the event study, the return of one single firm is not enough as returns are stochastic. Therefore, Fama et al. (1969) aggregates abnormal returns over many stocks and statistically test them to for the hypothesis that returns around the event date, are on average, not different from their expected returns.

An essential step for conducting an event study is identifying the event of interest and timing of the event. For this research, the event for the study is the publishment of the Newsweek green scores, and the timing is the announcement day of the particular list at the Newsweek's website. Newsweek uses financial and environmental data from a couple of years before the announcement, but this is not seen as the event date as the market did not have all this information at that time.

Furthermore, after establishing the event date, a proper estimation window, and the event window has to be chosen. The estimation window is created to calculate returns in periods before the event happens, to establish expected returns in the event period. Two factors are essential for choosing the estimation window. First, the window should be large enough to ensure that there are enough observations to estimate. Second, the window needs to be as close to the event date, but cannot be affected by the event itself. Furthermore, the event window should start at the moment where it fully captures the effect of the event. For this paper, an estimation window of six months before the announcement date until five trading days before the publishment is used. Six months are used because, in that way, the previous Newsweek Green Ranking does not affect the normal returns in the estimation window. Finally, an event window of three days prior to the announcement until three days after the announcement is chosen. This event window is further proved by MacKinlay (1997) as it states that it is customary to define the event window to be larger than the specific period of interest. Figure 3 illustrates the timeline around an event.

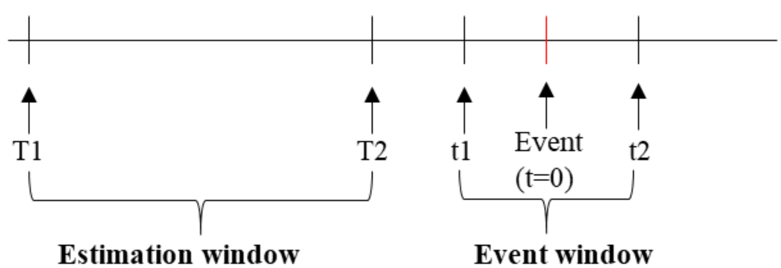


Figure 3. Timeline around an event

4.1.2 Benchmark model and cumulative abnormal returns

An import part of executing an event study is choosing the right benchmark model for the behavior of stocks. There are several models for calculating the behavior of stocks, but this study uses the market model as used by Fama (1969) to measure abnormal returns. For this model, the intercept and slope of a stock's return on the market return in the estimation window are used from regression to estimate stock's expected returns conditional on market returns during the even window. This model produces firm-specific expected return estimates instead of estimating expected return without constraining the cross-section of the returns. Formula 4.1 is used to calculate the daily returns by using the market model for company i at time t, and formula 4.2 calculates the normal daily returns in the event window through the estimates of the regression from formula 4.1. The country-specific sector index, as explained in section 3.2.2., is used for the market return.

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \tag{4.1}$$

$$NR_{it} = \hat{\alpha}_i + \hat{\beta}_i R_{mt} \tag{4.2}$$

Furthermore, after obtaining the normal daily returns from formula 4.2, the daily abnormal returns are easily calculated by subtracting the normal daily returns from the actual daily returns provided in the event window through the following formula.

$$AR_{it} = R_{it} - NR_{it} \tag{4.3}$$

In analyzing the formulas stated above, it is conventional to use time $t = 0$ as the event date. For formula 4.1, this means that all the periods indicated by t are negative as the estimation window occurs before the event.

This research, however, is interested in the cumulative abnormal returns (CAR) over the event window. It shows the cumulative effects of the wanderings of the returns on the effect of the event from their normal relationships to market movements (Fama, 1969). The usual way to compute the CAR is simply by summing up the daily abnormal returns over the event window as expressed by the following formula where t_0 is the start of the event window and t_1 is the end of the event window.

$$CAR_i = \sum_{t=t_0}^{t_1} AR_{it} \tag{4.4}$$

Since the CARs in the event study of this paper are aggregated over the cross-section, the cumulative average abnormal returns (CAAR) are used to calculate the t-statistic for the first hypothesis. The CAAR is with formula 4.5.

$$CAAR = \frac{1}{N}\sum_{i=0}^{N} CAR_i \qquad (4.5)$$

4.1.3 Testing cumulative abnormal returns

Overall, it is instructive to graphically report the cumulative abnormal returns. However, the analysis needs to be supported by statistical tests to test the first hypothesis made in section 2. The statistical tests are made to answer the question of whether the cumulative abnormal returns calculated for the top-ranked firms are significantly different from zero at a specific, prior specified, significance level. The null hypothesis tested in this paper for the first hypothesis is.

$$H_0: E(CAR_i) = 0.$$

Furthermore, to statistically test the null hypothesis, a couple of assumptions are made and discussed in this section. The first assumption is that the cumulative abnormal returns are independent and identically distributed, i.e., $Cov(CAR_{it}, CAR_{jt}) = 0, i \neq j$. The second assumption is that the cumulative abnormal returns follow a normal distribution with mean zero and variance σ^2. Unfortunately, the first assumption does not hold since one event affects multiple companies at the same time. Therefore, there may be a cross-sectional correlation between the CARs. The variance of the average cumulative abnormal returns is larger because it is no longer equal to $\frac{1}{N}$ multiplied to the variance of a single return. The usual t-statistic is therefore biased upwards, which concludes in the fact that the null hypothesis is rejected too often.

A good benchmark minimizes the problem of this cross-sectional dependence. Therefore, this paper uses 79 country-specific sector index returns to calculate the normal returns from formula 4.1 and 4.2. Newsweek provides the GICS sectors for all the companies on the Newsweek green score list, and these specific indexes are therefore used to minimize the cross-sectional dependence. Furthermore, the event study methodology automatically controls for heterogeneity by subtracting the normal returns from the actual returns.

Moreover, to test the first hypothesis of the positive returns around the announcement date of the Newsweek green scores, a t-test for the cumulative abnormal returns is calculated. First, the standard deviation of the CARs is calculated through formula 4.6.

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(CAR_i - CAAR)^2} \tag{4.6}$$

Finally, the cumulative average abnormal returns from formula 4.5 are divided with the standard deviation of formula 4.6. Hereafter, the square root of the total observations for the cumulative abnormal returns is multiplied, and this gives the t-statistic (TS) used for hypothesis one. The formula is used in the following way.

$$TS = \sqrt{N} * \frac{CAAR}{s} \sim N(0,1) \tag{4.7}$$

4.2 Environmental performance and firm value

In this section, the regression models used for the analysis of hypothesis two, and three are elaborated. First, the ordinary least squares (OLS) regression for the second hypotheses is explained. Second, the different effect on the overall stock price from firms from the European Union and the United States are further clarified.

4.2.1 Effect Newsweek Green score

To analyze whether the effect of the Newsweek green score list has a positive or negative effect on a firms' stock price, the cumulative abnormal returns from each company are regressed on the green scores and multiple control variables mentioned in section 3.3 by using ordinary least squares (OLS) estimations. The regression is done in the same way as expressed in the paper of Yadav et al. (2016) and constructed in the following form.

$$\begin{aligned} CAR_i = \alpha + \beta_1 GreenScore_i + \beta_2 ROA_i + \beta_3 GrowthRate_i + \beta_4 \ln(Sales)_i \\ + \beta_5 Leverage + \theta_i + \theta_t + \varepsilon_i \end{aligned} \tag{4.8}$$

Where CAR_i is the cumulative abnormal return for firm i from three days before the event until three days after the event. $GreenScore_i$ is the Newsweek green score for firm i, $\beta_2, \beta_3, \beta_4$, and β_5 are the regression coefficients of the control variables for each firm. Furthermore, θ_i presents the sector dummy for each firm and θ_t presents the year fixed effects for year t.

Furthermore, an additional analysis is done in this research to find the effect between companies which are continuously improving their environmental performance, enter the Newsweek green list for the first time, and when companies lower their environmental performance. By examining these results, additional information about the impact of the Newsweek green score list is collected.

4.2.2 International effect Green Score

Hypothesis 3 predicts that the cumulative abnormal returns around the event are more significantly related to a firms' green score in European Union firms as European Union firms disclose, on average, less environmental information and operate in a more intrinsic CSR economy. To test the third hypothesis, this paper uses three different regressions. First, the cumulative abnormal returns from companies of both the United States and the European Union are regressed on the green score and the respective control variables from formula 4.8. Furthermore, another regression uses an interaction term between the environmental green score and the firms' continent of origin, i.e. the European Union and the United States. An interaction term estimates the effect of one independent variable on the dependent variable on the magnitude of another dependent variable (Norton and Wang, 2004). For this reason, this paper runs an ordinary least squares regression to predict the cumulative abnormal returns as a function of the Newsweek green score, a dummy variable for European Union and the United States, and an interaction dummy between the two dependent variables. Furthermore, to finally test if the cumulative abnormal returns are more significantly related to the geographical location, the coefficient of the interaction dummy is investigated. The regressions used to test this third hypothesis is done in the following way.

$$CAR_i(EU) = \alpha + \beta_1 GreenScore_i + \beta_2 ROA_i + \beta_3 GrowthRate_i + \beta_4 \ln(Sales)_i + \beta_5 Leverage + \theta_i + \theta_t + \varepsilon_i \quad (4.9)$$

$$CAR_i(US) = \alpha + \beta_1 GreenScore_i + \beta_2 ROA_i + \beta_3 GrowthRate_i + \beta_4 \ln(Sales)_i + \beta_5 Leverage + \theta_i + \theta_t + \varepsilon_i \quad (4.10)$$

$$CAR_i = \alpha + \beta_1 GreenScore_i + \beta_2 Continent_i + \beta_3 (GreenScore_i * Continent_i) + \beta_2 ROA_i + \beta_3 GrowthRate_i + \beta_4 \ln(Sales)_i + \beta_5 Leverage + \theta_i + \theta_t + \varepsilon_i \quad (4.11)$$

Where $CAR_i(EU)$ are the cumulative abnormal returns for firms in the European Union firm i, $CAR_i(US)$ the CARs for firm i in the United States, and CAR_i the cumulative abnormal returns in the total sample. $Continent_i$ is a dummy variable which has the value of 1 if the firm on the Newsweek green score list is located in a European Union country and 0 if the headquarter of the company is located in the United States. Furthermore, $GreenScore_i * Continent_i$ is the interaction dummy between the green score and the continent where the headquarter of firm i is located.

4.3 Long-term mean reversion

This section explains the methodology used to test the last hypothesis on the long-term over- and underreaction to events. There are some differences between conducting a short-term event study and long-term event study, which are further explained in this section.

4.3.1 Benchmark model long-term event

In contrast to short-horizon event studies, the correction for market returns with the market model has several disadvantages in the long-horizon event studies. As stated in the paper by Fama and French (1993), there are well-known deviations from the capital asset pricing model, such as the book-to-market effect and the size effect. Therefore, the long-term event study is not corrected using only market returns as a benchmark, but with a model which takes into account the known factors which influences the stock price. Fama and French (1993) stated the first three known factors, as shown in the following formula.

$$R_{it} - R_{ft} = \alpha_i + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_{it} \tag{4.12}$$

In this formula the SMB_t is the ‘small minus big’ factor, which is the difference in return between a portfolio of small firms and the portfolio of large firms. Furthermore, HML_t is the ‘high minus low’ factor, and this is the difference between a portfolio of firms with a high book-to-market ratio and a portfolio with a low book-to-market ratio.

However, since the introduction of the Fama-French three-factor model, Fama and French (2015) introduced the Fama and French 5-factor model. In this model, the ‘robust-minus-weak’ operating profitability[13] factor (RMW_t) is taken into account. This factor is the difference between the return on diversified portfolios of stocks with robust and weak profitability. Furthermore, the ‘conservative-minus-aggressive’ investment factor takes into account the difference between the

[13] Formula for operating profitability: $OP = \frac{Revenue - COGS - Selling - administrative\ expenses - Interest\ expenses}{Book\ value\ of\ Equity}$

returns on diversified portfolios of the stocks of low and high investment[14] firms. The Fama and French (2015) five-factor is stated in the following formula.

$$R_{it} - R_{ft} = \alpha_i + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CMA_t + \beta_5 RMW_t + \varepsilon_{it} \qquad (4.13)$$

To generate the normal daily returns from this model in the event window, the estimates of the regression from formula 4.13 are used in the following way for the event window.

$$NR_{it} = R_{ft} + \hat{\beta}_1(R_{mt} - R_{ft}) + \widehat{\beta_2} SMB_t + \widehat{\beta_3} HML_t + \widehat{\beta_4} CMA_t + \widehat{\beta_5} RMW_t \qquad (4.14)$$

Furthermore, the daily abnormal returns are calculated in the same way, as stated in section 4.1.2. For the long-run, not the cumulative abnormal returns are used, but the buy-and-hold abnormal returns (BHAR). Barber & Lyon (1997) argue that long-run abnormal returns need to be calculated as the long-run buy-and-hold returns of a sample less the long-run return of an appropriate benchmark. The formal definition of the buy-and-hold abnormal return from Barber & Lyon (1997) is stated in the following way where H is the end period of the sample is days.

$$BHAR_i = \prod_{t=1}^{H}[1 + R_{it}] - \prod_{t=1}^{H}[1 + NR_{it}] \qquad (4.15)$$

Barber & Lyon (1997) state that CARs are biased predictors of BHARs, and this could lead to incorrect inferences. Furthermore, even if the CARs are correct, the magnitude does not correspond to the value of investing in the median or overage sample firm relative to the appropriate benchmark of the horizon. Finally, the t-statistics of BHARs are well-specified if the sample firms to control firms are similar in size and book-to-market ratios (Barber & Lyon, 1997). This paper controls for size and book-to-market ratios, as explained in the Fama-French three-factor model.

The calculations for the associated t-statistics have the same methodology as the variables explained in section 4.1.2 and are therefore not further explained in this section. Since this study only investigates a long-horizon period of nine months with only one event happening during that time in our analysis, the serial correlation between events from different months is not an issue.

[14] Formula for investment: $Inv_t = \frac{Asset_t - Asset_{t-1}}{Asset_t}$

4.3.2. Measuring long-term effect Green Score

For analyzing hypothesis four, both the buy-and-hold abnormal returns as well as the effect from the green score is investigated. First, the long-term buy-and-hold returns are regressed on the overall green score by using the same dependent variables as expressed in formula 4.8. However, unlike the event window of seven days, an event window of nine months is chosen to measure the long-term effect of the buy-and-hold returns. The regression is therefore done with the following formula:

$$\begin{aligned} BHAR_i = \alpha + \beta_1 GreenScore_i + \beta_2 ROA_i + \beta_3 GrowthRate_i + \beta_4 \ln(Sales)_i \\ + \beta_5 Leverage + \theta_i + \theta_t + \varepsilon_i \end{aligned} \tag{4.16}$$

In this regression, $BHAR_i$ is the long-term buy-and-hold abnormal return for the firm i over the nine-month sample period. To test if the long-term buy-and-hold returns are mean-reverting, a low and insignificant β_1 coefficient is expected.

Finally, to further check the model of Klassen and McLaughlin (1996) about the improvement of financial performance due to improved environmental performance, also the difference in buy-and-hold abnormal returns between the best 100 firms in the Newsweek green score list and the worst 100 first of the list are investigated for the nine-month sample period in each year of the sample. To execute this analysis, a different portfolio is made, which contains the 100 highest ranked environmental firms and a portfolio with the 100 lowest-ranked environmental firms on the Newsweek green score list. Both portfolios are analyzed individually and after that, subtracted from each other to analyze the differences. Formula 4.18 is used to calculate the t-statistic for the difference of the top 100 ranked firms i, and the bottom 100 ranked firms j.

$$TS = \frac{BHAR_i - BHAR_j}{\sqrt{\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}}} \tag{4.17}$$

4.4 Robustness

In order to ascertain the robustness of the findings in chapter 4, this paper implements four different robustness tests, focusing on the dependent variable, additional control cross-sectional correlation, a different subsample, different event date, and by randomizing the quantiles made for differences in buy-and-hold abnormal returns for the best- and worst 100 firms in the Newsweek green score.

Firstly, the assumption that all the variance of the abnormal returns is the same for all series i ($\sigma_i^2 = \sigma^2$) is not likely to be accurate as some stocks are more volatile than others. If there are some stocks which are more volatile than others in the analysis of the CARs in the short-term, then this could cause a more substantial variation in the average abnormal returns and could lead to a lower power of the t-statistic. Therefore, a weighted average of the abnormal returns is used to put a lower weight on abnormal returns with a high variance. This research uses a time-series estimate of the standard deviation of the abnormal returns in the estimation window.

Secondly, this paper uses 79 different benchmarks to correct for cross-sectional dependence. However, a cross-sectional correlation could still be a problem in the event, and therefore, an additional method is addressed to check the robustness. This method of cross-sectional correlation is used when assuming the correlation is not affected by the event itself, and the returns are serially uncorrelated. This method uses robust standard errors to correct for cross-sectional correlation. To apply this correction, first, the standard errors of the coefficients in the estimation window are estimated clustered by date. Hereafter, the coefficients in the event window are estimated, and this coefficient is divided by the square root of the total days in the event window times the total days in the estimation window and further multiplied by the standard error estimated in the first step.

Thirdly, a robustness test is executed by excluding the sample of 2014, 2015, and 2016. The methodology of these samples is slightly different from the 2017 methodology. The most important indicator, which has a total weight of 20% on the overall green score, is different in the 2014 sample. In 2014, the KPI was "reputation" instead of "Green Score," and this could, therefore, lead to a bias in the results. Moreover, a different numerical value is used in the sample of 2015 and 2016. Furthermore, the scores from the sample of 2017 contain two deduction variables which can deduct a maximum of five percent. Overall this could lead to biases in the overall results and to check the implications of this change; the robustness test will exclude the sample of 2014, 2015, and 2016.

Fourthly, to verify the robustness of our event study, three different simulation event dates are set one, three, and six months before the actual event. The same OLS regressions, as explained in section 4.2, are done on these event studies, and the significance of the green score is then further investigated. No significant statistical relationship should occur between the green score and the cumulative abnormal returns in these three event studies as this would imply that the announcement from Newsweek does not differ from any other date.

Furthermore, to check the robustness of the buy-and-hold abnormal returns for the 100 best firms and 100 worst firms on the Newsweek green score list, the sample of the quintiles is selected randomly, and the t-test is checked on significance. If the model of Klassen and McLaughlin (1996) is correct, there should be no statistically significant effect between the difference in abnormal returns over the nine-month sample period.

V. Results

This section provides the results of the empirical analysis of this paper. First, section 5.1 discusses the Pearson Correlation Matrix for the variables used in the analysis. Second, section 5.2 conducts the results on the first hypothesis about the effect of the Newsweek green score list. Third, section 5.3 reports the results of the second and third hypotheses about the different effect for European Union firms in comparison with United States' firms. Fourth, section 5.4 discusses the long-term mean reversion from the event. Finally, section 5.5 discusses the results from the additional robustness tests.

5.1 Pearson Correlation Matrix

The pairwise Pearson Correlation matrix is shown in Table A.2 of the Appendix between the green score, the key performance indicators, and the control variables. Since the key performance indicators are making up the overall green score, a strong positive correlation exhibits between the variables. However, the Green Revenue KPI is only slightly positively correlated to the overall green score and is even negatively correlated with most of the other KPI's. A possible explanation could be that in 2014, this score was a reputation score, which is changed in 2015 to a green revenue score. Since the reputation does not include actual environmental scores, this could lead to a lower correlation with the other variables. According to Table A.2, a higher growth rate leads to a lower green score. Furthermore, leverage correlates slightly negative with the overall green score. Additionally, the return on assets and the size of the firm are positively correlated with the green score. The positive correlation between return on assets and leverage could be explained by the paper of Jensen and Meckling (1976) as they state that higher leverage is associated with improved efficiency. By looking at the correlations, no control variable exceeds an absolute value 0.30, which indicates that no alarming correlations are in the dataset. All the correlations are significant on a 1% level except for the correlation between leverage and the green score.

5.2 Short-term analysis: Univariate analysis Green Score

The first hypothesis predicts an overall positive cumulative average abnormal for the 100 highest ranked firms of the Newsweek green score list. Table 2 presents the effect of the green ranking announcement on firm value for the 100 highest ranked firms for different event windows ranging from a three-day event window until a nine-day event window for the total sample, the United States, and the European Union. In the overall sample, a negative average cumulative abnormal return is shown in Table 2. However, all the results ranging from the event window are not significant, except for the three-day CAAR starting from $t = 0$. The CAAR (0, +2) therefore results in a negative average cumulative abnormal return of 0.17 percent from the event date until two days after the event date. Furthermore, only the three-day event window from United States' firms shows a negative, significant result in Table 2. On average, there is a negative cumulative abnormal return of 0.32 percent for the United States' companies after the announcement of the green score list. Firms from the European Union have a positive average cumulative abnormal return in the event window. However, these results are highly insignificant. Possible reasons for these insignificant results from Table 2 can be the low percentage of new companies entering the Newsweek green score list in the 100 highest ranked firms as 11%, 9%, and 10%, are new entrances for respectively the whole sample, the United States, and the European Union. These low percentages could lower the overall shock and therefore lower the cumulative abnormal returns of the total sample. According to the results of Table 2, the first hypothesis is not rejected.

Table 2

Univariate Analysis: Top 100

This Table presents the means in percentages and t-statistics of the cumulative average abnormal returns for the different event windows in parentheses for the 100 highest ranked firms on the Newsweek green score List for the whole sample (Overall), the United States (USA), and the European Union (EU). *, **, ***, indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively. The p-values are given in italic and between parentheses.

	Overall (N = 218)	**t-statistic**	**USA** (N=121)	**t-statistic**	**EU** (N=97)	**t-statistic**
CAAR (0, +2)	-0.17*	-1.73	-0.28	-1.59	-0.08	-0.77
	(0.09)		*(0.11)*		*(0.44)*	
CAAR (+1, -1)	-0.10	-1.08	-0.32*	-1.84	0.07	0.70
	(0.28)		*(0.07)*		*(0.49)*	
CAAR (+2, -2)	-0.04	-0.35	-0.19	-0.96	0.08	0.53
	(0.73)		*(0.34)*		*(0.60)*	
CAAR (+3, -3)	-0.14	-0.95	-0.35	-1.31	0.08	0.18
	(0.34)		*(0.19)*		*(0.86)*	

Furthermore, the univariate analysis over the total sample is stated in Table A.3 in the Appendix. This table presents an insignificant positive CAAR for the seven, and five days event window for the total sample as well as the European Union, and The United States subsample. The three-day event windows show insignificant negative CAARs for the total sample and subsamples, except for the three-day event window for the European Union. A possible explanation for the insignificant results is the different reactions from investors to low ranked companies in comparison to high ranked companies. Also, the different reaction to the environmental improvement or downturn of a company could lead to conflicting market reactions. This topic is further investigated in section 5.3 of this paper.

5.3 Short-term analysis: Multivariate analysis Green Score

Table 3 provides the OLS regressions of the second, and third hypothesis resulting in the effects between the cumulative abnormal returns and the green score from Newsweek. The first column presents the total sample without taking into account the industry fixed effects. The second column presents the total sample with the industry and year effects. The third until the fifth column presents OLS results for the European Union, The United States, and the total sample with the continent dummy- and interaction coefficient respectively. Moreover, the sixth column presents the score-up, new entrance, and score-down results. Finally, column seven presents the magnitude of the change in green scores between two contiguous years.

The first column of Table 3 expresses a statistically significant negative coefficient of the green score. Economically, this means that the announcement of the green score hurts the cumulative abnormal returns of the stock price for the whole sample. These results indicate that a one percentage point increase in the green score leads to a 1.1 basis point (bps) decrease in the cumulative abnormal return in the seven-day event window. This result is not in line with hypothesis 2 as the cumulative abnormal returns are not positively related to a firm's green score. However, this result does not take into account the industry fixed effects. This regression leads to an R^2 of 4.2% and an adjusted R^2 of 3.4%.

The OLS regression depicted in column (2) of Table 3 presents the effect of the announcement from Newsweek for the whole sample and by taking into account the industry fixed effects. As explained in section 3.3, the industry effect influences financial performance. However, after controlling for industry fixed effects, a negative and significant relationship still exists between the green score and the CARs. This outcome is not consistent with the theory of Yadav et al. (2016), and not in line with hypothesis 2. Comparing the results of column (2) with column (1), the impact of the green score is less harmful as an increase of 1% in the green score leads to a 1.0 bps decrease in the CAR. The R^2 of the regression in column 2 is higher in comparison with the regression from column 1, 5.0% and 4.2% respectively; however, the adjusted R^2 is lower by adding the industry fixed effects. The coefficient results of the green score indicate that investors believe an improved environmental performance leads to lower financial performance or are not willing to pay a premium for environmental friendly firms. This result is in line with the findings of Meric et al. (2012) as they also found an adverse reaction on the stock market and the green score.

The third until the fifth column presents the findings for the United States and the European Union subsample as well as the total sample with the continent dummy- and interaction coefficient between the green score and the geographical location, respectively. First, consistent with the overall findings of column (1) and (2), a negative and significant relationship is found for companies from the United States and the European Union. Overall, the announcement is more negatively related to the green score in European Union countries as a one percent increase in the green score leads to a 1.6 bps decrease in the CAR for European Union firms and 1.0 bps for United States firms. However, by looking at column (5), the different magnitude of the impact of the green score between European Union and the United States' companies is not significantly different from each other as seen by the interaction dummy coefficient from column (5), -0.6 bps. Furthermore, European Union firms have, on average, a 0.5 percent higher CAR. However, this result is also insignificant, and therefore, we can not conclude anything from this result. Due to this insignificant results, we can not reject hypothesis 3. Moreover, the adjusted R^2 is low for column (3) and (5) as it does not exceeds 3.5 percent. However, the adjusted R^2 is 13.8% for the European Union subsample and is, therefore, even higher than the adjusted R^2 from the results of Yadav et al. (2016). This concludes that the explanatory power from the variables of column (4) is high in comparison to other studies investigating the impact of Newsweek's green score (Meric et al., 2012; Yadav et al., 2016).

Table 3

Multivariate analysis short-term CAR

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily cumulative abnormal returns on the green score and multiple control variables displayed in the variables column. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column reports the results from the United States' subsample, European Union subsample, and the total sample with the dummy- and interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean-variance inflation factor, respectively. *, **, *** indicates the significance at the p<0.10, p<0.05, p<0.01 level respectively. The standard errors are given between parentheses.

Variables	**(1)** **No Dummy**	**(2)** **Total Sample**	**(3)** **USA**	**(4)** **Europe**	**(5)** **Europe-USA**	**(6)** **Improvement**	**(7)** **Δ**
Green Score	-0.011***	-0.010**	-0.010*	-0.016**	-0.010*	-0.013**	-0.008
	(0.004)	(0.004)	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)
ROA	-0.019	-0.023	-0.018	-0.027	-0.018	-0.019	0.004
	(0.015)	(0.016)	(0.020)	(0.029)	(0.016)	(0.020)	(0.020)
Growth	-0.006	-0.007	-0.013*	0.001	-0.006	-0.008	-0.016**
	(0.006)	(0.006)	(0.007)	(0.012)	(0.006)	(0.007)	(0.008)
Log(Size)	0.000	0.000	-0.000	0.001	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Leverage	-0.001***	-0.001***	-0.000	-0.004***	-0.001***	-0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
EU					0.005		
					(0.005)		
EU * Green Score					-0.006		
					(0.010)		
Dummy_New						0.003	
						(0.003)	
Dummy_Improvement						0.002	
						(0.003)	
Δ Green Score							-0.002
							(0.010)
Constant	0.000	0.005	0.007	0.004	0.004	0.019	0.007
	(0.009)	(0.009)	(0.012)	(0.015)	(0.009)	(0.013)	(0.012)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.042	0.050	0.042	0.182	0.052	0.051	0.026
Observations	885	885	569	316	885	655	513
Adjusted R-squared	0.034	0.032	0.015	0.138	0.032	0.026	-0.006
Mean VIF	1.23	1.41	1.46	1.43	2.26	1.44	1.40

To further investigate the effect of the green score, the sixth column presents the regression coefficients which includes dummies for firms which improved their green score and entered the Newsweek green score for the first time in comparison with companies which decreased their green score respectively. The coefficient of the green score is in line with the results obtained in the first five columns as the announcement on the Newsweek green score hurts the overall stock price. The dummy scores are in line with the theory of Klassen and McLaughlin (1996). In column (6) a positive coefficient of the dummy variable of firms which improved their green score and entered the green score for the first time are presented in comparison to firms which decreased their green score. Economically, this means that if a firm improves its green score, the cumulative abnormal return decreases by 0.2 bps higher in comparison with companies showing deterioration in their green score. Moreover, the CAR for firms which enter the Newsweek ranking for the firms time is 0.3 bps higher compared to firms which green score decreased over one year. However, the dummy variables are not significant and therefore, hard to interpret.

Moreover, the effect of the magnitude of the change in green score for companies on the green list, the Δ green score coefficient is included in the regression. This variable takes into account the magnitude of the change and investigates this effect. Column (7) in Table 3 presents the result and find a highly insignificant result of a decrease of 0.2 bps in the CAR for an increase of one percent of the green score in comparison to the green score of last year. Therefore, the magnitude of the increase has not a statistically significant effect on the cumulative abnormal returns. Finally, the results from column six and seven indicate that there is not a statistically significant difference between firms which improve their environmental performance and for firms which deteriorate their environmental performance.

Furthermore, additional analysis is done by looking at the relationship between the green score and the different quintiles of the whole sample and the two subsamples of the United States and the European Union. The results are displayed in Table A.4 in the Appendix. No statistically significant result is found by doing this analysis. Moreover, Table A.5 presents the relationship from all the individual KPI's with the cumulative abnormal returns. Only the audit score has a significant negative effect of 0.5 bps when a third party audited the firms on the cumulative returns for the whole sample, without including green revenue and reputation, and for the 2017 sample. The sustainable themed committee and the water productivity score are significant and positively related to the cumulative abnormal returns in 2014. Overall, there is not an obvious key performance indicator driving the cumulative abnormal returns.

Finally, the last row of Table 3 presents the results of the average variance inflation factor (VIF). The VIF test detects potential multicollinearity among the control variables. Multicollinearity arises when the control variables correlate with each other next to the correlation with the dependent variable resulting in a bias of the regression coefficient. According to O'brien (2007), the most common rule of thumb as a sign of severe multi-collinearity is a variance inflation factor of 10. Since the regressions show a maximum mean VIF of 2.26, it can be concluded that no multicollinearity problems exist in any regression in Table 3.

5.4 Long-term analysis: Effect of green scores

Table 4 presents the regression coefficient results for the fourth hypothesis about the long-term mean reversion of the abnormal results from the announcement. This section briefly explains the results obtained from the nine-month buy-and-hold abnormal returns.

Firstly, column (1) and (2) of Table 4 present a significant positive coefficient for the green score. Economically, this means that a 1% increase in the green score leads to a 14.9 bps (1) or 11.7 bps (2) increase in the nine-month buy-and-hold returns. With, and without including the industry fixed effects, the results are still significant on a 1% significance level. These results imply that higher environmental performance leads to better than expected financial performance. These results are in line with the model of Klassen and McLaughlin (1996). However, the low adjusted R^2 implies that the explanatory power of the variables in models (1) and (2) is small and the variables, therefore, have little impact on the daily buy-and-hold abnormal returns. The outcomes of the first two columns are not in line with hypothesis 4 as the BHAR are not mean-reverting but positively related to the green score in the long-term.

Secondly, by looking at the results from the United States subsample (3), European Union subsample (4), and the total sample with the continent- and interaction dummy (5), a positive green score, continent dummy, and interaction dummy coefficients are seen. However, no statistically significant conclusion can be obtained as the green score variable extremely insignificant as well as the continent dummy and the interaction dummy. Therefore, the results indicate that there is no

Table 4

Multivariate analysis of long-term BHAR

This Table presents the main results concerning the fourth hypothesis of this research by regressing the daily nine-months buy-and-hold abnormal returns on the green score and multiple control variables displayed in the variables column. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column report the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummies for the different quintiles are presented in column (6). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the $p<0.10$, $p<0.05$, $p<0.01$ level respectively. The standard errors are given between parentheses

Variables	(1) No Dummy	(2) Total Sample	(3) USA	(4) Europe	(5) Europe-USA	(6) Quintiles
Green Score	0.149***	0.117***	0.034	0.055	0.036	0.028
	(0.037)	(0.037)	(0.047)	(0.074)	(0.048)	(0.126)
ROA	-0.093	-0.234*	-0.132	-0.142	-0.082	-0.237*
	(0.127)	(0.134)	(0.154)	(0.278)	(0.136)	(0.134)
Growth	-0.050	-0.109**	-0.148***	-0.019	-0.108**	-0.096*
	(0.047)	(0.049)	(0.055)	(0.112)	(0.049)	(0.050)
Size	0.000	0.005	0.013	-0.012	0.005	0.004
	(0.007)	(0.007)	(0.009)	(0.013)	(0.007)	(0.007)
Leverage	0.001	-0.001	-0.001	0.000	0.000	-0.001
	(0.003)	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)
EU					0.062	
					(0.041)	
EU * Green Score					0.042	
					(0.079)	
Q2						0.011
						(0.028)
Q3						-0.002
						(0.041)
Q4						-0.025
						(0.056)
Q5						-0.063
						(0.077)
Constant	-0.035	-0.022	-0.115	0.229	-0.039	0.048
	(0.075)	(0.078)	(0.093)	(0.143)	(0.077)	(0.118)
Industry FE	No	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.026	0.060	0.077	0.089	0.087	0.067
Observations	877	877	563	314	877	877
Adjusted R-squared	0.017	0.043	0.050	0.040	0.067	0.046
Mean VIF	1.23	1.37	1.41	1.41	2.27	3.31

statistically significant difference between the performance of the Newsweek green score companies for the nine-month BHARs.

Finally, since the score-up, new entrance, and score-down coefficients are important for the short-term analysis as this could lead to a surprising effect in the short term event study, the actual environmental performance is more important to analyze in the long-term. Therefore, Table 4 presents also the regression results for the different quintiles of the green ranking. An insignificant coefficient of the green score is shown in column (6). Moreover, the dummy variables of the different quintiles show results, which are in line with the expectations since the quintiles show negative coefficients which are decreasing over the hight of the quintile. Only the dummy variable for the second quintile shows a positive coefficient, which implies that companies ranked in the second quintile show, on average, a 1.1 bps higher buy-and-hold abnormal returns over the nine-month holding period than companies in the first quintile. However, all the dummy coefficients for the quintiles are insignificant and are therefore hard to interpret economically.

To further analyze the theory of Klassen and McLaughlin (1996), Table 5 presents the nine-month buy-and-hold abnormal returns for the total sample, the 100 highest ranked firms, and the 100 lowest-ranked firms. There is a significant difference between the performance of the 100 best-ranked firms on the Newsweek green score list and the 100 bottom-ranked firms. Over nine months, the difference in the BHAR was 10.3%, with a 1% significance level. This BHARs are graphically presented in Figure 4 over the whole nine-months period, with also the whole sample included. The 100 highest ranked firms have an average nine-month BHAR of 3.4% with a standard deviation of 17.3% on a 1% significance level. Moreover, the 100 lowest-ranked firms generate a negative nine-month BHAR of 6.9% with a standard deviation of 27.1% on a 1% significance level. The overall sample performance on average a 0.8% BHAR but this results is highly insignificant with a t-statistic of 0.02, and therefore we can not conclude anything on the performance of the whole sample.

After analyzing the nine-month daily buy-and-hold abnormal returns for the whole sample, also the European Union and the United States' subsample is analyzed. Table A.6 and A.7, and graph A.1 and A.2 in the Appendix show the results for the univariate analysis for the European Union and the United States' subsample. For both samples, a negative nine-month daily BHAR is found; however, only the sample for the United States' firms is significant due to the low number of observations in the European Union sample. Moreover, for the European Union sample, a positive and significant BHAR is generated for the top 100 firms in the Newsweek Green Scor list where the top 100 firms in the United States' show insignificant and small negative nine-month BHARs.

Table 5

Long-term univariate analysis

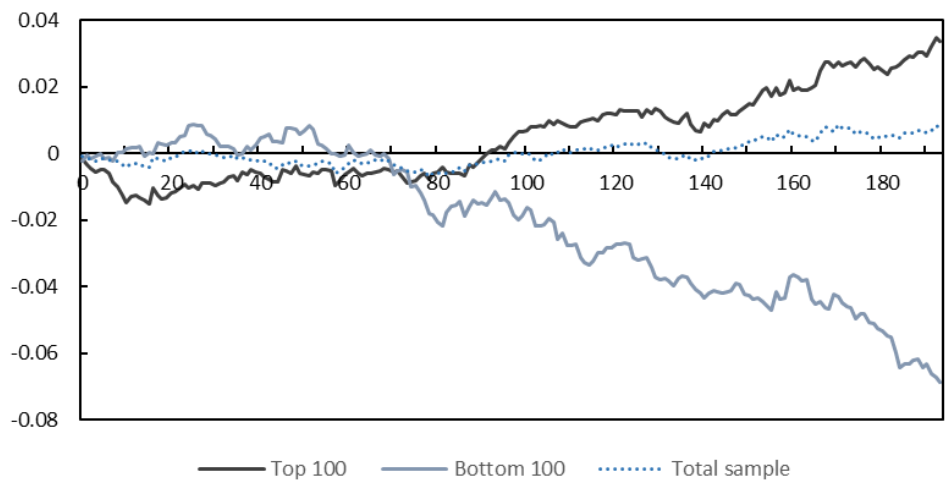
This Table presents the nine-months daily buy-and-hold abnormal returns for the total sample, top 100, and bottom 100 ranked firms, respectively. The last row shows the difference between the 100 highest- and lowest-ranked firms. The average BHARs, standard deviation, t-statistic, p-value, and total observations are presented in the columns, respectively. *, **, ***, indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively.

	BHAR	St. dev	t-statistic	p-value	Observations
Total Sample	0.008	0.214	0.02	0.981	895
Top 100	0.034***	0.173	2.87	0.005	217
Bottom 100	-0.069***	0.271	-2.76	0.007	118
Difference	0.103***	0.028	3.72	0.000	

Figure 4

Long-term univariate analysis

This Figure presents the nine-month daily buy-and-hold abnormal returns for the 100 highest-ranked firms, 100 lowest-ranked firms, and the whole sample. The vertical axes show the buy-and-hold abnormal performance as the horizontal shaft presents the trading days.



After explaining the long-term results, it can be concluded that there is evidence which partly supports the theory of Klassen and McLaughlin (1996) as they state that improved environmental performance leads to improved financial performance. However, this does not confirm hypothesis four as the results are not mean-reverting since there is a significant positive nine-month BHAR for the 100 best-ranked firms and a significant negative nine-month BHAR of the 100 bottom-ranked firms. Also, the regression coefficient is significantly positively related to the green score over the nine months for the whole sample.

5.5 Robustness

Firstly, to verify the results obtained in section 5.3, the cumulative abnormal returns are standardized to put a lower weight on the abnormal returns with a high variance. To obtain the standardized abnormal returns, the abnormal return is divided by the estimated standard deviation of the firm. The results are shown in Table A.8 in the Appendix. However, all the green score coefficients are insignificant, and the adjusted R^2 does not exceed 1.5% in any column. Therefore, the standardized cumulative abnormal returns are not in line with the results presented in section 5.3.

Secondly, The results to further control for cross-sectional correlation are displayed in table A.9 in the appendix. This method uses robust standard errors, as explained in section 4.4. No statistically significant results for the green score are found on the short-term analysis. This result is explained by the fact that the results obtained in section 5.3 are not highly significant, and due to the higher standard error, the results are insignificant. However, these results are not in line with the results in table 3 and therefore, do not confirm the main results of section 5.3.

Thirdly, the methodology of the Newsweek green score list of 2017 is slightly different from the one in 2014, 2015, and 2016. Therefore, an additional analysis is done by only taking the 2017 sample to verify the results obtained in section 5.3. The results are shown in Table A.5 in the Appendix. Table A.10 partly support the results obtained in section 5.3 as the total sample still shows a significant negative regression coefficient for all the green scores. However, the United States and the interaction column results are negative but insignificant. A possible explanation could be the limited amount of 111 and 119 observations, respectively. Striking is the adjusted R^2 of the European Union's subsample with an R^2 of 40.3%, which is extremely high. Finally, this robustness test mostly verifies the results obtained in section 5.3.

Fourthly, the results of the multivariate regression for the different event dates are shown in table A.11, A.12, and A.13 in the Appendix. Almost all the results are insignificant and fluctuate between positive and negative green score coefficients. This result is in line with the expectation as insignificant, and various results are expected for event studies without an actual event occurring.

Finally, to ascertain the results obtained in section 5.4, the first and fifth quintile are selected randomly. These results are shown in Table A.14 and Figure A.3 in the Appendix. An insignificant negative result is found for the top quintile and an insignificant positive result for the fifth quintile. These outcomes verify the obtained results from section 5.4 as a random quintile does not generate significant daily abnormal returns, and these findings are in line with the theory of Klassen and McLaughlin (1996).

VI. Conclusion

This chapter presents the results, limitations, and recommendations for future research. First, section 6.1 discusses the main findings and what is the main takeaway from the paper. Second, section 6.2 addresses the limitations of this research and ascertains the recommendations for future research.

6.1 Discussion

This study aims to investigate the impact of environmental performance on financial performance in an international environment by using the short- and long-term event study methodology on the announcement of the Newsweek green scores in 2014 until 2017. Firstly, this research investigated the short-term impact of the announcement from Newsweek for the 100 highest ranked firms. Secondly, the relationship between the green score and the cumulative abnormal returns is investigated for the whole sample and additionally by investigating the difference between firms from the European Union and United States' companies. Finally, the long-term effect of the announcement is investigated.

First, this study shows that no market reaction occurs for the 100 highest ranked firms for European Union companies and United States firms when Newsweek announces their green rankings. These findings are contrary to the findings of Amato & Amato (2012) as they found a positive impact on stock prices for firms ranked in the top quartile of the list. There are multiple explanations for this finding. First, the surprising effect for investors could be decreasing due to the low percentage of new entrance firms into the 100 highest ranked firms. Second, the social influence of traditional newsweeklies could be waning, and therefore, the impact on the stock market could be decreasing. A concern when estimating the short-term reaction of an event that affects multiple companies is a cross-sectional dependency. To address this concern, this study uses 79 country- and sector-specific benchmarks, which solves most of the cross-sectional dependence.

Second, by regressing the seven-day CAR on the green scores and multiple control- and dummy variables, a negative short-term relationship between the CARs and the green score is found for both European Union companies and United States firms. Also, a negative relationship between the green score and the CARs is found when controlled for improvement, new entrance and deterioration in the green score in comparison to the year before. This result implies that higher environmental performance leads to a negative stock market performance in the short-term. These results are contrary to the findings of Yadav et al. (2016) since they found a positive short-term relationship for the short-term green score coefficient. However, the findings are in line with the results of Meric et al. (2012). Moreover, this study finds no different market reaction between companies headquartered in the European Union and firms headquartered in the United States. This result does not confirm the third hypothesis. These findings imply that lower environmental disclosure for European countries (Bartolomeo, 2000; Thomas et al., 1997) and the different environmental perception for European countries (Matten and Moon, 2008) does not lead to a higher surprising effect for European Union companies on the stock market when Newsweek releases their green score list. However, the increasing global economic and financial integration could be a possible explanation for these results.

Third, this paper uses the daily nine-month buy-and-hold abnormal returns with the event study methodology to investigate the long-term effect of the announcement from Newsweek with the Fama-French 5-factor model as a benchmark. The results partly imply that high environmental performance leads to improved financial performance in the long-term for European Union firms and United States companies. Moreover, companies ranked in the top quintile with their environmental performance perform significantly better than firms ranked in the bottom quintile of the green ranking over the nine months. However, other factors than the green score are driving this performance since no significant effect is found for each quintile. Moreover, these results partly support the theory of Klassen and McLaughlin (1996), who argued that improved environmental performance leads to improved financial performance. However, this result does not confirm hypothesis four since the results are not mean-reverting in the long-term.

Although the results do not confirm any of the four hypotheses, the results do provide evidence concerning the main research question by showing that the Newsweek green score leads to an adverse stock market reaction in the short-term and that the green score is positively related to the long-term stock market performance. Especially the difference between the 100 highest-ranked firms and the 100 bottom-ranked firms show a highly significant difference in the financial performance in the stock market in the long-term. Moreover, this research does not find any differences in reaction to the stock market between firms from the European Union and companies from the United States.

Overall, the main takeaway of this research indicates that it 'pays to be green' in the long-term for large multinationals in the United States and the European Union. Meaning that firms are rewarded for having excellent environmental performance. The results contain valuable information for listed companies and their environmental practices since environmental performance is a significant determinant of investment decisions.

6.2 Limitations and future research

Firstly, this research uses data from the 500 largest companies based on revenue and market capitalization. Therefore, only large firms are investigated in this research. However, studies found that the size of a firm has a significant effect on the degree of proactiveness since larger organizations are more likely to adapt to proactive environmental practices (Buysse and Verbeke, 2003; Russo and Fouts, 1997). Although this research uses size as a control variable, the total sample only contains large firms, and therefore, there is sample bias. Moreover, further research could extend the sample to include small- and midsized companies to exclude the bias which exists by only including large companies. However, since the green ranking of Newsweek contains multiple key environmental indicators and does not contain small- and midsized companies, this research only investigated large companies.

Secondly, this research only contains European Union firms and companies from the United States. Saida (2009) states that both the United States and the EU have the most rigid environmental regulations in the world. Moreover, there are also different environmental rules and environmental disclosures across continents. For this reason, the research is only applicable to companies in the EU and the United States, and this is an essential implication of the research. Further research could extend the sample to more geographically located companies to provide a more accurate estimate of the impact of environmental performance on financial performance. However, since the different effect of the announcement of the green scores of Newsweek is never investigated between continents, this research started by comparing the European Union and the United States.

Thirdly, the literature to measure environmental performance is extensive. Prior literature has used a wide range of indicators to measure environmental performance since it is hard to measure every environmental aspect of a firm. This research uses data from Newsweek, which consists of eight different environmental indicators. However, other papers use different environmental indicators or environmental awards to measure environmental performance. For this reason, a limitation of this study is the measurement of environmental performance by 'only' eight environmental indicators. Future research could investigate the effect of environmental performance with more advanced environmental indicators.

References

Accenture. & UNGC. (2010). A new era of sustainability: UN Global Compact-Accenture CEO study 2010. New York: United Nations Global Compact and Accenture

Adams, C. A., & Kuasirikun, N. (2000). A comparative analysis of corporate reporting on ethical issues by UK and German chemical and pharmaceutical companies. *European Accounting Review*, *9*(1), 53-79.

Aerts, W., Cormier, D., & Magnan, M. (2008). Corporate environmental disclosure, financial markets and the media: An international perspective. *Ecological Economics*, *64*(3), 643-659.

Amato, L. H., & Amato, C. H. (2012). Environmental policy, rankings and stock values. *Business Strategy and the Environment*, *21*(5), 317-325.

Bandyopadhyay, S. P., Brown, L. D., & Richardson, G. D. (1995). Analysts' use of earnings forecasts in predicting stock returns: Forecast horizon effects. *International Journal of Forecasting*, *11*(3), 429-445.

Barber, B. M., & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of financial economics*, *43*(3), 341-372.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, *49*(3), 307-343.

Bartolomeo, M., Bennett, M., Bouma, J. J., Heydkamp, P., James, P., & Wolters, T. (2000). Environmental management accounting in Europe: current practice and future potential. *European Accounting Review*, *9*(1), 31-52.

Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of financial economics*, *8*(3), 205-258.

Buysse, K., & Verbeke, A. (2003). Proactive environmental strategies: A stakeholder management perspective. *Strategic management journal*, *24*(5), 453-470.

Cañón-de-Francia, J., & Garcés-Ayerbe, C. (2009). ISO 14001 environmental certification: a sign valued by the market?. *Environmental and Resource Economics*, *44*(2), 245-262.

Capon, N., Farley, J. U., & Hoenig, S. 1990. Determinants of financial performance: A meta-analysis. Management Science, 36: 1143-1159.

Carroll, A. B. (1979). A three-dimensional conceptual model of corporate performance. *Academy of management review*, *4*(4), 497-505.

Chynoweth, E., & Kirshner, E. (1993). Environmental standards provide competitive advantage. *Chemical Week*, *152*(16), 46-52.

Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, *63*(4), 1977-2011.

Cormier, D., Magnan, M., & Van Velthoven, B. (2001). Environmental reporting management: an international perspective. *Document de travail- Centre de recherche en gestion*.

Daniel, K. D., Hirshleifer, D. A., & Subrahmanyam, A. (1997). A theory of overconfidence, self-attribution, and security market under-and over-reactions. *Self-Attribution, and Security Market Under-and Over-Reactions (February 19, 1997)*.

De Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, *40*(3), 793-805.

de La Bachelerie, V., & Boillet-Mongodin, J. (1993). Le traitement de l'environnement dans les rapports annuels. *Revue Française de Comptabilité*, (246), 79-85.

Edwards, W. (1968). Conservatism in human information processing. *Formal representation of human judgment*.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, *49*(3), 283-306.

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

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3-56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, *116*(1), 1-22.

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International economic review*, *10*(1), 1-21.

Flammer, C. (2013). Corporate social responsibility and shareholder reaction: The environmental awareness of investors. *Academy of Management Journal*, *56*(3), 758-781.

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

Guthrie, J., & Parker, L. D. (1990). Corporate social disclosure practice: a comparative international analysis. *Advances in public interest accounting*, *3*, 159-175.

Ilinitch, A. Y., Soderstrom, N. S., & Thomas, T. E. (1998). Measuring corporate environmental performance. Journal of accounting and public policy, 17(4-5), 383-408.

International Organization for Standardization (ISO). (1999). ISO 14031: 1999. Environmental management—environmental performance evaluation—guidelines. Geneva: ISO.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, *3*(4), 305-360.

Judge, W. Q., & Douglas, T. J. (1998). Performance implications of incorporating natural environmental issues into the strategic planning process: An empirical assessment. *Journal of management Studies*, *35*(2), 241-262.

Kahneman, D. & Tversky, A. (1982). "Intuitive Prediction: Biases and Corrective Procedures." In D. Kahneman, P. Slovic, and A. Tversky, (eds.), Judgment Under Uncertainty: Heuristics and Biases. London: Cambridge University Press, 1982.

Keele, D. M., & DeHart, S. (2011). Partners of USEPA climate leaders: an event study on stock performance. *Business Strategy and the Environment*, *20*(8), 485-497.

Klassen, R. D., & McLaughlin, C. P. (1996). The impact of environmental management on firm performance. *Management Science*, *42*(8), 1199-1214.

Luo, X., & Bhattacharya, C. B. (2006). Corporate social responsibility, customer satisfaction, and market value. *Journal of marketing*, *70*(4), 1-18.

Lyon, T. P., & Shimshack, J. P. (2015). Environmental disclosure: Evidence from Newsweek's green companies rankings. *Business & Society*, *54*(5), 632-675.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of economic literature*, *35*(1), 13-39.

Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383-417.

Mathur, L. K., & Mathur, I. (2000). An analysis of the wealth effects of green marketing strategies. *Journal of Business Research*, *50*(2), 193-200.

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.

Meric, I., Watson, C. D., & Meric, G. (2012). Company green score and stock price. *International Research Journal of Finance and Economics*, *82*(1), 15-23.

Mitchell, M. L., & Stafford, E. (2000). Managerial decisions and long-term stock price performance. *The Journal of Business*, *73*(3), 287-329.

Newey, W. K., & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 777-787.

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, *4*(2), 154-167.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, *41*(5), 673-690.

Opler, T. C., & Titman, S. (1994). Financial distress and corporate performance. *The Journal of finance*, *49*(3), 1015-1040.

Porter, M., & Van der Linde, C. (1995). Green and competitive: ending the stalemate. *The Dynamics of the eco-efficient economy: environmental regulation and competitive advantage*, *33*.

Roe, M. J. (2003). *Political determinants of corporate governance: Political context, corporate impact*. Oxford University Press on Demand.

Rosewicz, B. (1990). Americans are willing to sacrifice to reduce pollution, they say. *Wall Street Journal*, *20*, A1

Ruf, B. M., Muralidhar, K., Brown, R. M., Janney, J. J., & Paul, K. (2001). An empirical investigation of the relationship between change in corporate social performance and financial performance: A stakeholder theory perspective. *Journal of business ethics*, *32*(2), 143-156.

Russo, M. V., & Fouts, P. A. (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of Management Journal*, *40*(3), 534-559.

Saida, D. (2009). Contribution on the analysis of the environmental disclosure: a comparative study of American and European multinationals. *Social Responsibility Journal*, *5*(1), 83-93.

The New York Times (2007). “Executive on a Mission: Saving the Planet.” Retrieved from http://www.nytimes.com/2007/05/22/science/earth/22ander.html

Thomas, P. B., & Kenny, S. Y. (1997, April). Environmental reporting: a comparison of annual report disclosures and popular financial press commentary. In *Interdisciplinary Perspectives on Accounting Practice conference, Manchester, England.[on-line]. Retrieved April* (Vol. 4, p. 2000).

Trump, C., Endrikat, J., Zopf, C., & Guenther, E. (2015). Definition, conceptualization, and measurement of corporate environmental performance: A critical examination of a multidimensional construct. Journal of Business Ethics, 126(2), 185-204.

Whitley, R. (1999). *Divergent capitalisms: The social structuring and change of business systems*. OUP Oxford.

Yadav, P. L., Han, S. H., & Rho, J. J. (2016). Impact of environmental performance on firm value for sustainable investment: Evidence from large US firms. *Business strategy and the environment*, *25*(6), 402-420.

Zeng, S. X., Meng, X. H., Yin, H. T., Tam, C. M., & Sun, L. (2010). Impact of cleaner production on business performance. *Journal of Cleaner Production*, *18*(10-11), 975-983.

Appendix

Table A. 1

These Tables report the distributions of the observations by countries (Table A.1.1), by sector (Table A.1.2), and by year (Table A.1.3) over the period 2014-2017. Sectors are classified using the Global Industry Classification Standard (GICS).

Table A.1.1: Sample distribution across countries.

Countries	Number of Observations	Percentage of total sample (%)	Average Green Score	Average cumulative abnormal return(%)
European	319	35.25%	0.515	0.195%
Austria	1	0.11%	0.301	0.159%
Belgium	3	0.33%	0.261	-0.149%
Denmark	4	0.44%	0.653	0.546%
Finland	5	0.55%	0.550	-0.430%
France	80	8.84%	0.531	0.117%
Germany	73	8.07%	0.484	0.152%
Ireland	4	0.44%	0.696	0.006%
Italy	14	1.55%	0.382	-0.484%
Luxembourg	1	0.11%	0.281	4.533%
Netherlands	21	2.32%	0.508	1.530%
Norway	3	0.33%	0.611	-0.113%
Spain	16	1.77%	0.535	0.012%
Sweden	16	1.77%	0.610	-0.175%
United Kingdom	78	8.62%	0.525	0.193%
United States	586	64.75%	0.387	0.030%
Total	**905**	**100.00%**	**0.432**	**0.088%**

Table A.1.2: Sample distribution across countries

Sectors	Number of Observations	Percentage of total sample (%)	Average Green Score	Average cumulative abnormal return(%)
Consumer Discretionary	172	19.01%	0.396	0.429%
Consumer Staples	130	14.36%	0.437	0.067%
Energy	89	9.83%	0.327	0.270%
Health Care	132	14.59%	0.436	-0.031%
Industrials	138	15.25%	0.463	0.207%
Information Technology	100	11.05%	0.444	-0.188%
Materials	46	5.08%	0.496	-0.455%
Telecommunication Services	37	4.09%	0.536	-0.195%
Utilities	61	6.74%	0.470	-0.071%
Total	**905**	**100.00%**	**0.432**	**0.088%**

Table A.1.3: Sample distribution across years.

Year	Number of Observations	Percentage of total sample (%)	Average Green Score	Average cumulative abnormal return(%)
2014	236	26.08%	0.407	-0.311%
2015	239	26.41%	0.462	-0.056%
2016	230	25.41%	0.476	0.112%
2017	200	22.10%	0.375	0.703%
Total	**905**	**100.00%**	**0.432**	**0.088%**

Table A. 2

Cross-Correlation of Variables

This Table presents the pairwise Pearson correlations below the diagonal for the periods 2014-2017. The Green Revenue variable only contains the years 2014-2016 since the variable had a percentage range in 2017. *, **, *** indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively. The numbers in the upper row represent the transposed variables as indicated on the vertical.

	1	2	3	4	5	6	7	8	9	10	11	12	13
Green Score	1.00												
Energy	0.73***	1.00											
Greenhouse Gas	0.73***	0.69***	1.00										
Water	0.73***	0.59***	0.50***	1.00									
Waste	0.69***	0.45***	0.42***	0.53**	1.00								
Pay Link	0.57***	0.21***	0.23***	0.29***	0.31***	1.00							
Board	0.55***	0.30**	0.33***	0.33***	0.33***	0.44***	1.00						
Audit	0.65***	0.39**	0.36***	0.42***	0.42***	0.54***	0.47***	1.00					
Green Revenue	0.19***	-0.02***	0.04***	-0.09***	-0.08***	-0.07***	-0.07***	-0.12***	1.00				
ROA	0.11***	0.10***	-0.12***	0.09***	0.09***	0.05***	0.01***	0.02***	0.03***	1.00			
Growth	-0.12***	-0.05***	-0.07***	-0.13***	-0.21***	-0.21***	-0.19***	-0.16***	0.26***	0.10***	1.00		
Size	0.11***	0.09***	0.15***	0.07***	0.19***	0.19***	0.20***	0.17***	-0.23***	-0.03***	-0.21***	1.00	
Leverage	-0.02***	-0.04***	-0.02***	-0.02***	0.03***	0.03***	-0.05***	0.01***	-0.01	0.02***	-0.01**	-0.00	1.00

Table A. 3

Univariate Analysis: Total Sample

This Table presents the means in percentages and t-statistics of the cumulative average abnormal returns for the different event windows in parentheses on the Newsweek green score List for the whole sample (Overall), the United States (USA), and the European Union (EU). *, **, ***, indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively. The p-values are given in italic and between parentheses.

	Overall (N=905)	**t-statistic**	**USA** (N=586)	**t-statistic**	**EU** (N=319)	**t-statistic**
CAAR (0, +2)	-0.05	-1.05	-0.05	-0.75	-0.06	-0.79
	(0.29)		*(0.45)*		*(0.43)*	
CAAR (+1, -1)	-0.01	-0.25	-0.08	-1.08	0.10	1.22
	(0.80)		*(0.28)*		*(0.22)*	
CAAR (+2, -2)	0.08	1.18	0.05	0.52	0.15	1.36
	(0.24)		*(0.60)*		*(0.17)*	
CAAR (+3, -3)	0.09	1.05	0.03	0.28	0.20	1.45
	(0.29)		*(0.78)*		*(0.15)*	

Table A. 4
Multivariate Analysis short-term Quintiles

This Table presents the main results concerning the green score of this research by regressing the daily cumulative abnormal returns on the green score, different quintile dummies, and multiple control variables displayed in the variables column. The first until the third column reports the results for the total sample, US subsample, and EU subsample respectively. The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the p<0.10, p<0.05, p<0.01 level respectively. The standard errors are given between parentheses.

Variables	(1) Q - Total	(2) Q - US	(3) Q - EU
Green Score	-0.010	-0.003	-0.008
	(0.015)	(0.020)	(0.022)
ROA	-0.023	-0.031	-0.018
	(0.016)	(0.029)	(0.020)
Growth	-0.007	0.002	-0.013*
	(0.006)	(0.012)	(0.007)
Log(Size)	0.000	0.001	-0.000
	(0.001)	(0.001)	(0.001)
Leverage	-0.001***	-0.004***	-0.000
	(0.000)	(0.001)	(0.000)
Q2	-0.000	-0.002	0.002
	(0.003)	(0.005)	(0.005)
Q3	-0.000	0.001	0.002
	(0.005)	(0.007)	(0.007)
Q4	-0.001	0.005	0.001
	(0.007)	(0.010)	(0.009)
Q5	0.000	0.014	0.002
	(0.009)	(0.014)	(0.013)
Constant	0.005	-0.007	0.005
	(0.014)	(0.021)	(0.019)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
R-squared	0.050	0.189	0.043
Observations	885	316	569
Adjusted R-squared	0.028	0.134	0.008
Mean VIF	3.33	2.37	4.13

Table A. 5
Multivariate Analysis short-term KPI's

This Table presents the main results concerning the key performance indicators from the green sore of this research by regressing the daily cumulative abnormal returns on the green score, KPI's, and multiple control variables displayed in the variables column. The first regression reports the results for the whole sample without the reputation and green score. The second column reports the results from 2014 including the reputation score. The third and fourth column presents the results including the green score of the 2015-2016, and 2017 sample respectively. The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the $p<0.10$, $p<0.05$, $p<0.01$ level respectively. The standard errors are given between parentheses.

Variables	(1) Total Sample	(2) 2014	(3) 2015-2016	(4) 2017
Green Score	-0.005	-0.494	2.795	-0.019
	(0.017)	(0.304)	(1.847)	(0.032)
Energy	-0.006	0.074	-0.424	-0.016
	(0.005)	(0.047)	(0.277)	(0.011)
Carbon	0.002	0.073	-0.418	0.014
	(0.005)	(0.047)	(0.277)	(0.013)
Water	0.003	0.081*	-0.417	0.001
	(0.004)	(0.046)	(0.277)	(0.010)
Waste	0.002	0.069	-0.418	0.015
	(0.004)	(0.046)	(0.277)	(0.010)
Pay Link	0.002	0.046	-0.275	0.004
	(0.003)	(0.030)	(0.185)	(0.007)
Themed Committee	-0.000	0.027*	-0.144	0.002
	(0.002)	(0.015)	(0.092)	(0.006)
Audit Score	-0.005**	0.023	-0.145	-0.016**
	(0.003)	(0.015)	(0.092)	(0.007)
ROA	-0.027*	-0.031	-0.030	0.015
	(0.016)	(0.029)	(0.022)	(0.044)
Growth	-0.007	-0.004	-0.016**	0.006
	(0.006)	(0.012)	(0.008)	(0.014)
Log(Size)	0.000	0.001	-0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.003)
Leverage	-0.001***	-0.001	-0.000	-0.003***
	(0.000)	(0.001)	(0.000)	(0.001)
Reputation		0.091		
		(0.061)		
Green Score (2015/2016)			-0.552	
			(0.369)	
Green Score (2017)				0.055
				(0.035)
Constant	0.006	-0.002	0.016	-0.011
	(0.010)	(0.016)	(0.014)	(0.032)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R-squared	0.058	0.062	0.055	0.173
Observations	885	230	456	199
Adjusted R-squared	0.032	-0.033	0.007	0.075
Mean VIF	2.42	122.02	5704.83	2.35

Table A. 6

Long-term univariate analysis for the European Union

This Table presents the nine-month daily buy-and-hold abnormal return, standard deviation, t-statistic, p-value, and observations for the total sample, the 100 highest-ranked firms, and the 100 lowest-ranked firms on the Newsweek green score list for the European Union. *, **, ***, indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively.

	BHAR	St. dev	t-statistic	p-value	Observations
Total Sample	0.068***	0.221	5.47	0.000	317
Top 100	0.061***	0.170	3.51	0.001	97
Bottom 100	-0.055	0.237	-0.81	0.434	12
Difference	0.116	0.070	1.64	0.127	

Figure A. 1

Long-term univariate analysis for the European Union

This Figure presents the European Unions nine-month daily buy-and-hold abnormal returns for the 100 highest-ranked firms, 100 lowest-ranked firms, and the whole sample. The vertical axes show the buy-and-hold abnormal performance as the horizontal shaft presents the trading days.

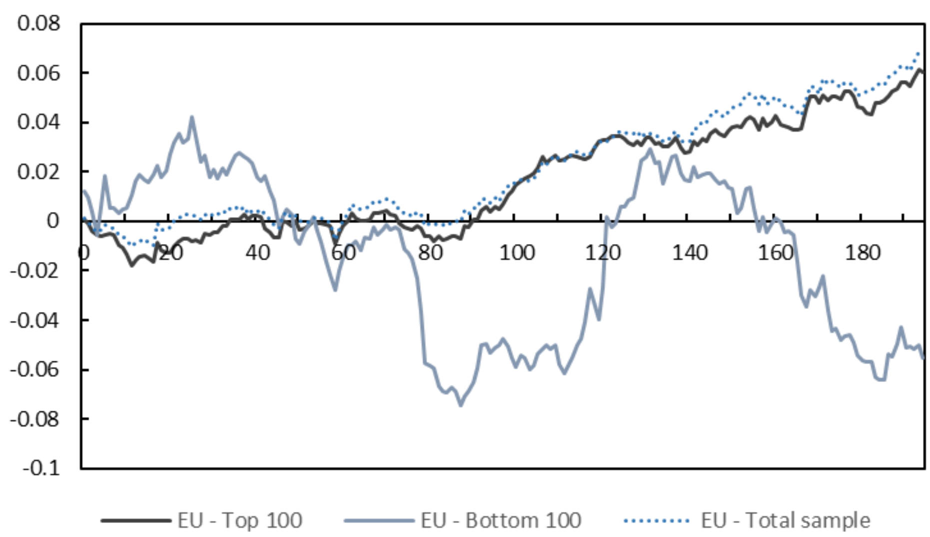


Table A. 7

Long-term univariate analysis the United States

This Table presents the nine-month daily buy-and-hold abnormal return, standard deviation, t-statistic, p-value, and observations for the total sample, the 100 highest-ranked firms, and the 100 lowest-ranked firms on the Newsweek green score list for the United States. *, **, ***, indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively.

	BHAR	St. dev	t-statistic	p-value	Observations
Total Sample	-0.025***	0.202	-3.01	0.003	578
Top 100	-0.001	0.170	-0.07	0.944	120
Bottom 100	-0.070***	0.275	-2.64	0.010	106
Difference	0.069**	0.031	2.24	0.027	

Figure A. 2

Long-term univariate analysis for the United States

This Figure presents the United States nine-month daily buy-and-hold abnormal returns for the 100 highest-ranked firms, 100 lowest-ranked firms, and the whole sample. The vertical axes show the buy-and-hold abnormal performance as the horizontal shaft presents the trading days.

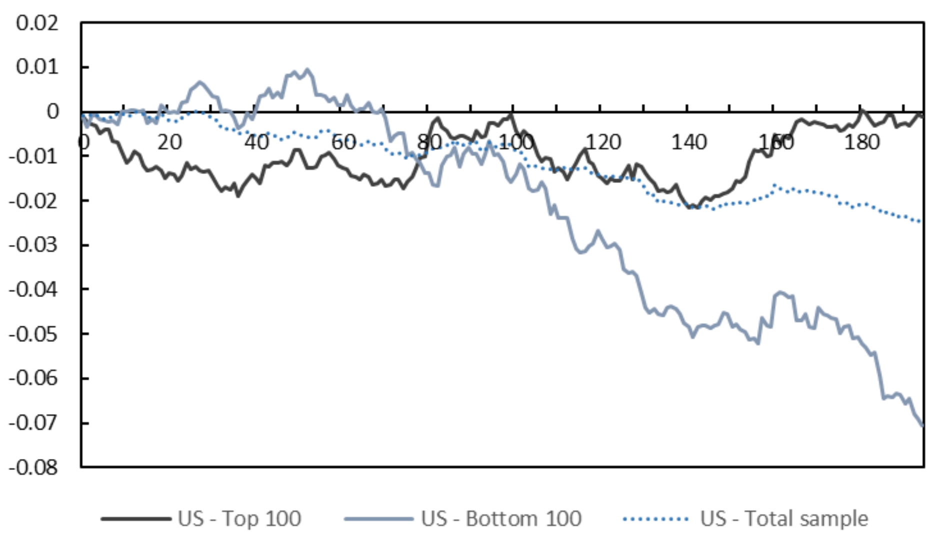


Table A. 8

Multivariate analysis short-term SCAR

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily standardized cumulative abnormal returns on the green score and multiple control variables displayed in the variables column. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column reports the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the $p<0.10$, $p<0.05$, $p<0.01$ level respectively. The standard errors are given between parentheses.

Variables	(1) No Dummy	(2) Total Sample	(3) USA	(4) Europe	(5) Europe-USA	(6) Improvement	(7) Δ
Green Score	-0.212	-0.164	-0.448	-0.421	-0.466	-0.311	-0.269
	(0.394)	(0.407)	(0.485)	(0.904)	(0.522)	(0.499)	(0.661)
ROA	-1.665	-0.817	-0.893	-0.507	-0.558	-0.577	-0.285
	(1.374)	(1.454)	(1.617)	(3.375)	(1.496)	(1.740)	(2.141)
Growth	-0.652	-0.688	-0.651	-0.642	-0.715	-0.934	-1.330
	(0.509)	(0.536)	(0.571)	(1.362)	(0.540)	(0.624)	(0.824)
Log(Size)	0.030	-0.034	-0.040	-0.034	-0.032	-0.090	-0.054
	(0.077)	(0.080)	(0.093)	(0.160)	(0.080)	(0.101)	(0.121)
Leverage	-0.030	-0.032	-0.036	0.005	-0.032	-0.033	-0.047
	(0.030)	(0.031)	(0.032)	(0.078)	(0.031)	(0.033)	(0.045)
EU					0.523		
					(0.875)		
EU * Green Score					-0.094		
					(0.456)		
Dummy_New						-0.015	
						(0.267)	
Dummy_Improvement						0.352	
						(0.224)	
Δ Green Score							0.740
							(1.121)
Constant	-0.249	0.460	0.628	0.502	0.472	1.435	1.339
	(0.805)	(0.845)	(0.970)	(1.734)	(0.852)	(1.132)	(1.385)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.010	0.029	0.021	0.065	0.030	0.032	0.036
Observations	885	885	569	316	885	655	517
Adjusted R-squared	0.001	0.011	-0.007	0.014	0.010	0.006	0.005
Mean VIF	1.23	1.37	1.41	1.41	2.27	1.44	1.47

Table A. 9

Multivariate analysis short-term Robust CAR

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily robust cumulative abnormal returns on the green score and multiple control variables displayed in the variables column. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column report the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the $p<0.10$, $p<0.05$, $p<0.01$ level respectively. The standard errors are given between parentheses.

Variables	**(1) No Dummy**	**(2) Total Sample**	**(3) USA**	**(4) Europe**	**(5) Europe-USA**	**(6) Improvement**	**(7) Δ**
Green Score	-0.011	-0.010	-0.010	-0.016	-0.01	-0.013	-0.008
	(0.007)	(0.007)	(0.009)	(0.014)	(0.009)	(0.009)	(0.011)
ROA	-0.019	-0.023	-0.018	-0.027	-0.018	-0.019	0.004
	(0.034)	(0.037)	(0.048)	(0.042)	(0.038)	(0.028)	(0.041)
Growth	-0.006	-0.007	-0.013	0.001	-0.006	-0.008	-0.016
	(0.014)	(0.015)	(0.018)	(0.020)	(0.015)	(0.011)	(0.016)
Log(Size)	0.000	0.000	-0.000	0.001	0.000	-0.000	0.000
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Leverage	-0.001*	-0.001*	-0.000	-0.004***	-0.001*	-0.001*	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EU					0.005		
					(0.009)		
EU * Green Score					-0.006		
					(0.016)		
Dummy_New						0.003	
						(0.005)	
Dummy_Improvement						0.002	
						(0.004)	
Δ Green Score							-0.002
							(0.017)
Constant	0.000	0.005	0.007	0.004	0.004	0.019	0.007
	(0.020)	(0.021)	(0.029)	(0.026)	(0.021)	(0.017)	(0.026)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Improvement FE	No	No	No	No	No	Yes	No
R-squared	0.042	0.050	0.042	0.182	0.052	0.051	0.026
Observations	885	885	569	316	885	655	513
Adjusted R-squared	0.034	0.032	0.015	0.138	0.032	0.026	-0.006
Mean VIF	1.23	1.41	1.46	1.43	2.26	1.44	1.40

TILBURG UNIVERSITY

Table A. 10

Multivariate analysis short-term CAR subsample

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily cumulative abnormal returns on the green score and multiple control variables displayed in the variables column for 2017. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column report the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the p<0.10, p<0.05, p<0.01 level respectively. The standard errors are given between parentheses.

Variables	**(1) No Dummy**	**(2) Total Sample**	**(3) USA**	**(4) Europe**	**(5) Europe-USA**	**(6) Improvement**	**(9) Δ**
Green Score	-0.018*	-0.019*	-0.016	-0.027*	-0.008	-0.014	-0.010
	(0.009)	(0.010)	(0.013)	(0.016)	(0.014)	(0.011)	(0.011)
ROA	0.021	0.014	0.067	0.035	0.021	0.032	0.086**
	(0.039)	(0.043)	(0.052)	(0.082)	(0.048)	(0.046)	(0.039)
Growth	0.005	0.005	-0.024	0.017	0.009	0.008	0.002
	(0.013)	(0.015)	(0.018)	(0.025)	(0.015)	(0.015)	(0.017)
Log(Size)	0.001	0.002	0.002	0.003	0.002	0.003	0.003
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Leverage	-0.002***	-0.002***	0.000	-0.008***	-0.002***	-0.002***	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
EU					-0.026		
					(0.021)		
EU * Green Score					0.011		
					(0.010)		
Dummy_New						0.006	
						(0.005)	
Dummy_Improvemer						-0.004	
						(0.006)	
Δ Green Score							0.002
							(0.019)
Constant	0.007	-0.004	-0.014	-0.015	-0.009	-0.020	-0.036
	(0.029)	(0.031)	(0.038)	(0.048)	(0.032)	(0.034)	(0.031)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Improvement FE	No	No	No	No	No	Yes	No
R-squared	0.082	0.113	0.112	0.492	0.120	0.122	0.101
Observations	199	199	111	88	199	199	124
Adjusted R-squared	0.058	0.050	-0.008	0.403	0.048	-0.015	-0.015
Mean VIF	1.03	1.35	1.42	1.52	2.08	1.51	1.51

Table A. 11

Multivariate analysis short-term CAR t=-1

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily cumulative abnormal returns on the green score and multiple control variables displayed in the variables column for the event window one month earlier. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column report the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the $p<0.10$, $p<0.05$, $p<0.01$ level respectively. The standard errors are given between parentheses.

Variables	(1) No Dummy	(2) Total Sample	(3) USA	(4) Europe	(5) Europe-USA	(6) Improvement	(7) Δ
Green Score	0.011*	0.010	0.013	-0.001	0.010	0.015**	0.007
	(0.006)	(0.006)	(0.008)	(0.011)	(0.008)	(0.008)	(0.007)
ROA	0.036*	0.041*	0.020	0.105**	0.040*	0.006	-0.019
	(0.021)	(0.022)	(0.028)	(0.041)	(0.023)	(0.026)	(0.023)
Growth	0.003	0.005	0.007	0.001	0.005	0.017*	0.026***
	(0.008)	(0.008)	(0.010)	(0.016)	(0.008)	(0.009)	(0.009)
Log(Size)	-0.003**	-0.003**	-0.004***	-0.000	-0.003**	-0.003**	-0.002
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
Leverage	0.001**	0.001**	-0.000	0.005***	0.001**	0.001**	-0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
EU					-0.001		
					(0.013)		
EU * Green Score					0.000		
					(0.007)		
Dummy_New						0.000	
						(0.004)	
Dummy_Improvement						-0.003	
						(0.003)	
Δ Green Score							-0.006
							(0.012)
Constant	0.021*	0.015	0.039**	-0.013	0.015	0.012	0.016
	(0.012)	(0.013)	(0.017)	(0.021)	(0.013)	(0.017)	(0.015)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.025	0.038	0.045	0.162	0.038	0.059	0.073
Observations	887	887	571	316	887	657	517
Adjusted R-squared	0.017	0.020	0.018	0.117	0.018	0.034	0.044
Mean VIF	1.23	1.37	1.41	1.41	2.27	1.43	1.47

Table A. 12

Multivariate analysis short-term CAR t=-3

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily cumulative abnormal returns on the green score and multiple control variables displayed in the variables column for the event window three months earlier. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column report the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the p<0.10, p<0.05, p<0.01 level respectively. The standard errors are given between parentheses.

Variables	(1) No Dummy	(2) Total Sample	(3) USA	(4) Europe	(5) Europe-USA	(6) Improvement	(7) Δ
Green Score	-0.001	0.001	0.003	0.005	0.001	0.001	-0.001
	(0.005)	(0.005)	(0.007)	(0.010)	(0.007)	(0.006)	(0.008)
ROA	-0.018	-0.033*	-0.025	-0.069*	-0.040**	-0.063***	-0.044*
	(0.017)	(0.018)	(0.022)	(0.035)	(0.019)	(0.022)	(0.026)
Growth	-0.004	-0.008	-0.021**	0.024*	-0.009	-0.011	-0.011
	(0.006)	(0.007)	(0.008)	(0.014)	(0.007)	(0.008)	(0.010)
Log(Size)	0.003***	0.003***	0.001	0.007***	0.003***	0.002*	0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Leverage	0.001**	0.001**	0.001*	0.001*	0.001**	0.001*	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
EU					0.010		
					(0.011)		
EU * Green Score					-0.008		
					(0.006)		
Dummy_New						-0.004	
						(0.003)	
Dummy_Improvemer						-0.002	
						(0.003)	
Δ Green Score							-0.019
							(0.014)
Constant	-0.025**	-0.020*	0.002	-0.066***	-0.018*	-0.005	-0.002
	(0.010)	(0.011)	(0.014)	(0.018)	(0.011)	(0.014)	(0.017)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.026	0.050	0.062	0.116	0.053	0.076	0.067
Observations	886	886	571	315	886	656	517
Adjusted R-squared	0.017	0.033	0.035	0.069	0.034	0.051	0.038
Mean VIF	1.23	1.37	1.41	1.41	2.27	1.40	1.47

Table A. 13

Multivariate analysis short-term CAR t=-6

This Table presents the main results concerning the second and third hypothesis of this research by regressing the daily cumulative abnormal returns on the green score and multiple control variables displayed in the variables column for the event window six months earlier. The first two regressions report the results for the whole sample without (1), and with (2) the industry fixed effects. The third until the fifth column report the results from the United States' sample, European Union sample, and the total sample with the interaction term for the continents, respectively. The results of the dummy for improvement, new entrance, and deterioration of the green score are presented in column (6). Lastly, the result of the difference in green score between two consecutive years is displayed in column (7). The last four rows present the R^2, total observations, adjusted R^2, and mean variance inflation factor, respectively. *, **, *** indicates the significance at the p<0.10, p<0.05, p<0.01 level respectively. The standard errors are given between parentheses.

Variables	**(1) No Dummy**	**(2) Total Sample**	**(3) USA**	**(4) Europe**	**(5) Europe-USA**	**(6) Improvement**	**(7) Δ**
Green Score	-0.001	-0.002	0.006	-0.010	0.007	-0.001	-0.009
	(0.005)	(0.005)	(0.007)	(0.008)	(0.006)	(0.006)	(0.007)
ROA	0.031*	0.032*	0.017	0.048*	0.026	0.030	0.061***
	(0.017)	(0.018)	(0.024)	(0.029)	(0.018)	(0.022)	(0.022)
Growth	-0.009	-0.009	-0.004	-0.012	-0.008	-0.009	-0.009
	(0.006)	(0.007)	(0.009)	(0.012)	(0.007)	(0.008)	(0.009)
Log(Size)	0.001	0.001	0.003**	-0.000	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Leverage	0.000	0.000	-0.000	0.002**	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
EU					-0.017		
					(0.011)		
EU * Green Score					0.004		
					(0.006)		
Dummy_New						-0.005	
						(0.003)	
Dummy_Improvemer						-0.004	
						(0.003)	
Δ Green Score							-0.009
							(0.012)
Constant	-0.016*	-0.020**	-0.036**	0.002	-0.021**	-0.013	-0.016
	(0.010)	(0.010)	(0.014)	(0.015)	(0.010)	(0.014)	(0.015)
Industry FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.018	0.023	0.031	0.082	0.030	0.030	0.045
Observations	884	884	569	315	884	655	517
Adjusted R-squared	0.009	0.005	0.003	0.033	0.010	0.004	0.014
Mean VIF	1.23	1.37	1.40	1.41	2.27	1.43	1.48

Table A. 14

Randomized long-term univariate analysis

This Table presents the nine-month daily buy-and-hold abnormal return, standard deviation, t-statistic, p-value, and observations for the randomly selected 100 highest-ranked firms, and the randomly selected 100 lowest-ranked firms on the Newsweek green score list for the European Union. *, **, ***, indicate significance at the p<0.10, p<0.05, and p<0.01 level respectively.

	BHAR	St. dev	t-statistic	p-value	Observations
Top 100	-0.006	0.227	-0.421	0.675	163
Bottom 100	0.009	0.201	-0.458	0.646	203
Difference	-0.015	0.023	-0.664	0.961	

Figure A. 3

Randomized long-term univariate analysis

This Figure presents the nine-month daily buy-and-hold abnormal returns for the randomly selected 100 highest-ranked firms and the randomly selected 100 lowest-ranked firms. The vertical axes show the buy-and-hold abnormal performance as the horizontal shaft presents the trading days.