universität wien

MASTERARBEIT

Titel der Masterarbeit

"Sustainability and Performance:
An Empirical Study on Sustainability Ratings
and Defining Sustainable Performance Drivers"

Verfasst von

Richard Geza Boulanger BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2015

Studienkennzahl lt. Studienblatt:	A 066 915
Studienrichtung lt: Studienblatt:	Masterstudium Betriebswirtschaft
Betreuer:	Univ.-Prof. Dr. Gyöngyi Lóránth

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Wien, am 15.01.2015

_____

Richard Geza Boulanger

Table of Contents

Table of Figures

List of Variables and Parameter

α_i = abnormal risk-adjusted return of an risky asset

α_p = abnormal risk-adjusted return of the portfolio

β_{HML} = exposure to value firms

β_m = systematic risk exposure to the market portfolio

β_{MOM} = exposure to momentum

β_{SMB} = exposure to small firms

ε_t = standard error at time t

HML_t = return through value firms at time t

MOM_t = return through momentum firms at time t

MV = firms market value

n = number of observations

P_t = stock price at time t

q = Tobins Q

SMB_t = return through small firms at time t

$SP_{p,t}$ = portfolios sharp ratio at time t

r_{3M} = 3 month EURIBOR return p.a.

$r_{f,t}$ = risk-free rate at time t

$r_{f,week}$ = weekly risk-free rate

$r_{m,t}$ = market return at time t

$r_{p,t}$ = portfolio return at time t

$\sigma_{i,M}$ = covariance between the asset return and market return

σ_M^2 = variance of the market return

σ_p = standard deviation of portfolio returns

V_T = value of firms tangible assets

V_I = value of firms intangible assets

1. Introduction

Socially Responsible Investments (SRI), which nowadays are often just called responsible investments (RI) (Benson and Humphrey, 2007), changed from a niche to almost a mainstream financial component. This type of investment approach is now associated with one of the megatrends in this millennium (Aburdene, 2012). Assets under management (AuM) did not only increase in absolute and relative terms since the first SRI fund was launched in 1971, growth of the sector was even much stronger than the general investment universe (USSIF, 2010). Today Socially Responsible Investments have passed $10 trillion. More precisely, the industry accounts nowadays for approximately 22% of AuM, which stands for a market value of $13.6 trillion worldwide. This includes all kinds of investments with a focus on environmental, social, and governance (ESG) criteria (GSIA, 2013). Most of the assets under management in the SRI sector are located in developed markets. 96% of all investments are located in Canada, Europe, and the USA (GSIA, 2013). Asia only accounts for about 0.6% in responsible investments. Interestingly, both private and public investors, who became more aware of the long-term influence of their investments, drove growth in Socially Responsible Investments (Fung et al., 2010). The segment gained importance just after the subprime crisis and beside the investors' awareness, academics interests were driven into the Socially Responsible Investments direction. In the last years the number of publications in this field increased exponentially.

Through the increasing interest in Socially Responsible Investments a debate about its profitability emerged, so many academics focused on the relationship between social and financial performance (Barnet and Solomon, 2006). On the one hand, academics believe that environmental, social, and governance performances lead to a competitive advantage and therefore could and should bring new investment opportunities (Derwall et al., 2005). If this idea holds, it could be said that higher SRI ratings lead to a financial outperformance. For this to be efficiently true, the market would have to price such ratings accordingly. So, if the market is efficient and the Efficient Market Hypothesis (EMH) holds, the market would absorb the better SRI rating instantly and no further outperformance could be generated. Another thought on SRI ratings and outperformance, which is related to an old financial theory, is the diversification effect. Since responsible investments use an exclusionary[1] or a best-in-

[1] Exclusionary approach: Investors eliminate certain industries out of the investment universe (e.g. Military)

class[2] approach, the investor is not able to invest in all companies, which should result in inferior risk-adjusted returns (Renneboog et al., 2008a). Furthermore, critics argue that investment allocations aimed to improve social and environmental performance in a company would cause respective product prices to rise and ultimately shareholder value to fall (Derwall et al., 2005). According to financial theory, this would conflict with companies' fundamental goal to increase shareholders value (Hiller et al., 2010). In contrast to the just mentioned critic, academics argue that responsible firms attract better and higher skilled management (Renneboog et al., 2008a). This handpicked management would enable the firm to generate long-term outperformance.

As already mentioned above, all these pros and cons for sustainable investments were generated through an increasing interest in the academic environment, which led to an increasing number of studies in the sustainability field. A various number of these studies, with a focus on mutual funds, such as Bauer et al. (2005, 2006, 2007), Renneborg et al. (2008b), Cortez, Silva, and Areal (2009), Galema, Plantinga, and Scholtens (2008), find no statistically significant results. Most of the studies focused on North American countries, but even the study from Renneborg at al. (2008b) with a global focus was not able to find overall significant results for performance and stated that no performance difference can be clearly observed. Some other studies in the sustainability field, such as Lee, Faff and Langfield-Smith (2007) examine an underperformance of sustainable investments in a selective range of countries. In contrast to these findings, Derwall et al. (2005) and Kempf and Osthoff (2007) find in their selected portfolios a statistically significant outperformance, but such findings, as mentioned before, are not in consensus with most of the sustainability studies. However, it's worth to mention that all studies are based on SRI fund performance, this directs these studies towards a measurement of the fund manager and his stock picking ability (Derwall et al., 2005), and so they lag from a quantitative database. SRI Index performance studies may overcome such empirical problems, but other problems will arise. By using SRI-Indices the problem of different SRI definitions emerge. In France for example nuclear power plants are seen as sustainable, since no CO_2 emission is produced. In contrast to France, the Austrian citizens do not see nuclear power as sustainable because of the nuclear waste.

[2] Best-in-Class: Only invest in the most sustainable companies out of each industry.

To eliminate the empirical weaknesses, our central empirical question will use a quantitative database to eliminate the stock picking bias and with no definition how to define sustainability, at least in accordance to exclusion criteria. Furthermore, we will not focus on any specific sample region, so we receive an overall picture. Our research is based on data from the research agency "oekom research AG", over the period from December 2005 until December 2013. oekom research AG is a sustainable research agency, based in Germany, which certified its research approach by ARISTA. By looking at the data it seems that on a regional level, there is a clear tilt towards the European region. This result could be explained by the fact that the European SRI market is by far the largest and most sophisticated one (GSIA, 2013). To generate a convincing and stable result, this study is based on the longest time horizon possible and we analyze different samples to identify if sustainability performed different in different market environments. To identify the performance of different sustainability levels, we constructed portfolios out of the most sustainable and least sustainable firms. All portfolios are constructed on a value-weighted basis and are compared to the Morgan Stanley Capital International World Index (MSCI World Index). We use three different approaches to measure performance along the study. First we use a head-to-head analysis, second we use the Capital Asset Pricing Model to identify if there is any significant out- or underperformance and last but not least, we use the Fama and French three-factor model (1993) to see if the performance comes from sustainability, small firm effect, growth effect, or just from the market. This research methodology goes hand in hand with most of the research in the sustainability field. With our approach we should be able to eliminate the regional bias, the stock picking bias and the SRI definition problem.

Furthermore, none of the studies mentioned above actually looks at the question why sustainability should outperform or which factors can be seen as performance driver or performance inhibitor. Academics mainly focus on the critics like diversification and costs, but none of them looks closely into the environmental, social, or corporate governance component (E, S, and G). Hence, we decided to go into more detail after our empirical study and look at literature, which analyzes performance drivers in the environmental, social, and corporate governance dimension. Corporate Governance is defined as the strongest performance driver in Socially Responsible Investments. This is based on the fact that corporate governance reduces the dissipation of money in a business (Dittmar and Mahrt-Smith, 2007). In other words, it eliminates part of the agency problem. Dittmar et al. (2007) states that

well governed firms are less affected by empire building and thus more value is transferred to shareholders. In the social dimension from Socially Responsible Investments, it is not that easy to identify if it is a driver or inhibitor, since the effects are difficult to measure (Edmans, 2011). On the one side the company faces costs to increase employees satisfaction and on the other side a satisfied employee has a higher motivation to work and will be more productive (Akerlof and Yellen, 1986). A motivated worker should produce more value for shareholders, but the question arises if the costs are higher than the benefits or the other way around. The answer to the question if the environmental dimension is a performance driver or inhibitor has changed from a doubtful picture to a clear one. Scholars state that environmental investments are a win-win situation for both, the investor and the firm (Konar and Cohen, 2001). In fact environmental investments are costly, but afterwards they will bring cost reductions and increase revenue (Jacobs, Singhal, and Subramanian, 2010). Furthermore, firms with a higher environmental level are expected to lower their reputation risks.

In our study we will search, through a literature review, for performance drivers and inhibitors from ESG, but we will not analyze the interdependencies between the different environmental, social, and governance factors.

Overall this study is structured as follows. Section one will provide an in-depth explanation where the idea of responsible investments comes from, how the development occurred, how sustainability can be measured, and which components are relevant. In section two we will go into our empirical study. Here, the methodology, database, and our portfolio formation will be explained. Section three will contain the results of the three different measurements. In section 4 we will give a literature review, concerning the fact why firms with a good environmental, social or governance behavior can lead to an under- or outperformance. It will be split in three parts to give a good overview of each dimension from E, S, and G. After all, there will be a wrap up in section 5 with a short conclusion.

2. Sustainability and its development

Sustainability stands for long-term thinking. The focus lies on resources, stability, and natural regeneration. The idea is that today's generation does not negatively influence the well-being of the next generation. So companies can behave in a sustainable way. But were they always interested in doing so and who is affected by their sustainable behavior?

2.1 Business Ethics

How corporations should act by doing business has changed over time. In one of his known articles „The Social Responsibility of Business Is to Increase Its Profits", Milton Friedman (1970) argues: "*The one and only social responsibility of business is to use resources and engage in activities designed to increase its profits*". Since the society has realized that profits always come with cost, and not uncommonly for the whole society, Friedmans attitude can no longer be maintained from today's perspective.

The so-called Pyramid of Corporate Social Responsibility from Carroll (1991) shows the principle existing responsibilities of today's economy and their rankings.

Graphic 1: Pyramid of Corporate Social Responsibility

Source: Carroll (1991), page 42

In addition to economic and legal responsibilities, which a company has to follow to survive, a corporation can/will also face and follow ethical and philanthropic responsibilities. Nowadays, from the perspective of sustainable investments, companies should at least accept and comply with three types of responsibility - economic, legal, and ethical. At least, by looking at the development of sustainable investments, this is what the society seems to demand. The more a company fulfills its economic and legal responsibility or even "over-fulfills" them, the better the CSR-Ranking will be and responsible investors will be more attracted to the corporation.

2.2 Affected Stakeholders by ESG

The four pillars from Carroll (1991) affect all stakeholders of a company and further, a part of the society. Stakeholders of a company are those parties who can formulate a claim against the company. These claims can be based on the basis of an economic relationship, company participation or on the basis of environment relations to the company.

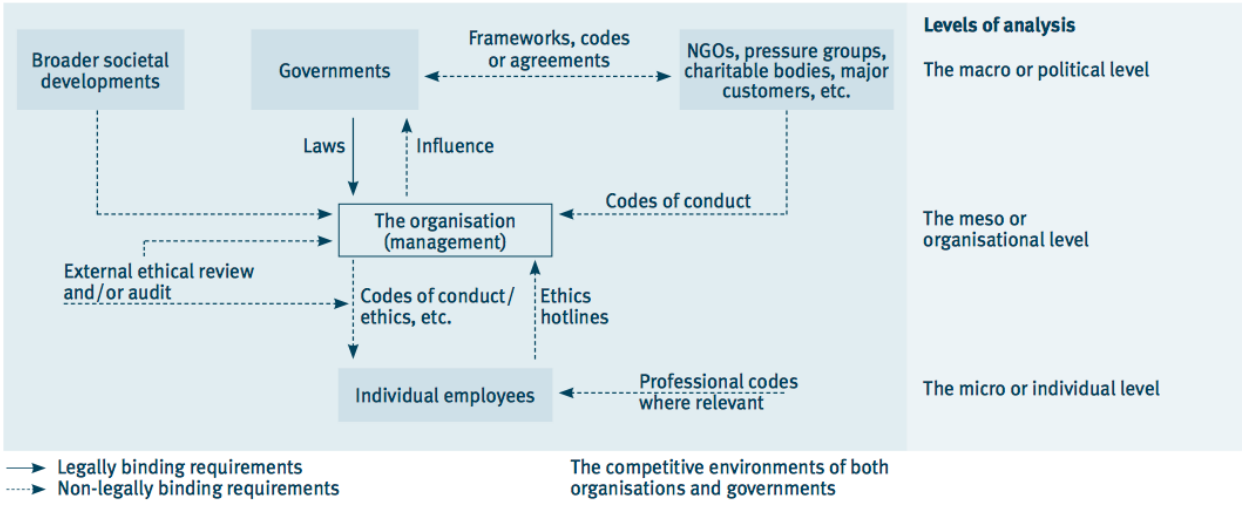
Stakeholder-related issues are usually also of ethical concern because they include obligations towards primary- or essential stakeholder groups. Consumers, shareholders, employees, and society are probably the most important stakeholders of any organization. Trevino and Nelson (2007) argue that there is nothing, which is more important to a company, then *“people who pay to have goods made, people who make them, people who buy them, and the place where the goods are made”* (Trevino and Nelson, 2007).

Rappaport is often seen as the "inventor" of the shareholder value concept with a clear focus on the economic view. Nevertheless, he argued that the investor should seek for those strategies that bring sustainable competitive advantage for the business. The concept of Rappaport says that a business can only survive over a long period if and only if relations with stakeholders are taken into account and the financial demands of stakeholders are met (Pinner, 2008).

The evaluation of stakeholder relations and therefore the valuation of extra-financial items is the main component of SRI ratings. So SRI/ESG research agencies assess the value of stakeholder relations to identify the potential of a sustainable future development.

These days corporations are constantly under pressure from "the world surrounding them" to act in a sustainable and ethical way. Here, pressure can occur from politics, organizations or individuals. The idea behind the political and organizational level is straightforward; under the individual level we can cluster employees, investors, clients, and supplier. Different stakeholders will interfere on different tasks and different intensities. To receive a comprehensive understanding of different linkages between the stakeholders and the possibilities to put ethical and sustainable pressure on a corporation, graphic 2 has been added.

Graphic 2: Formal and Informal linkage between stakeholders



Source: Fisher & Lovell (2006), page 379

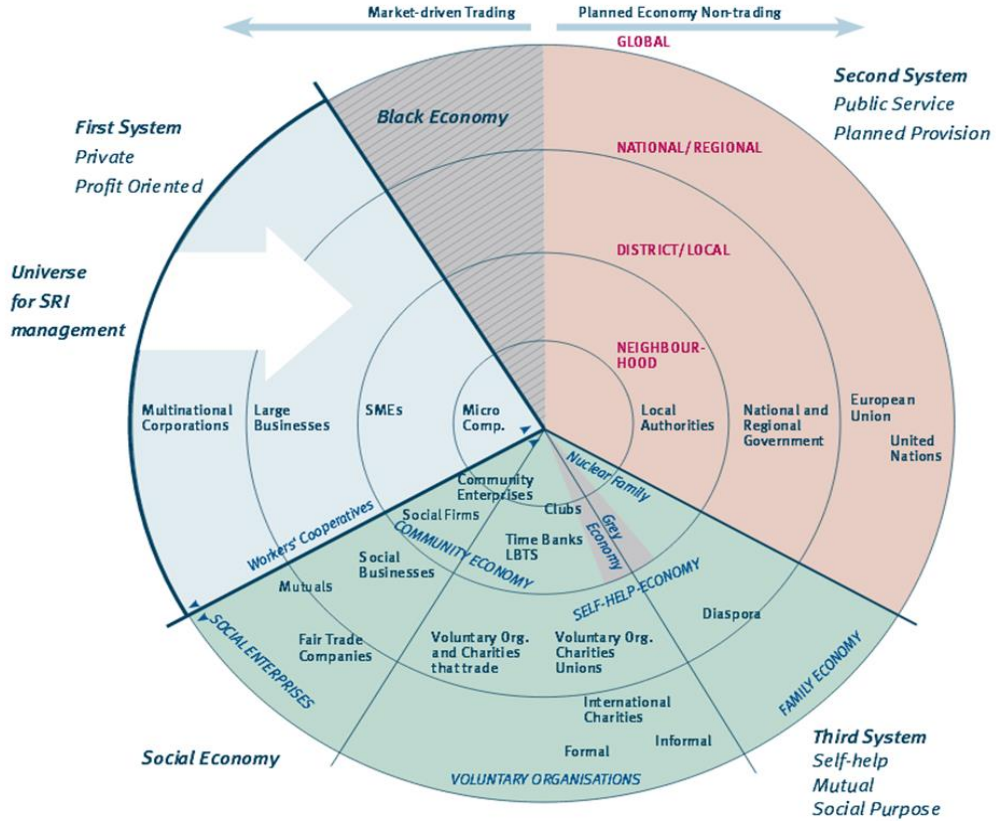
As it can be seen, firms' sustainability influences every stakeholder and every stakeholder is able to influence the company. Over all it can be said, that sustainability is relevant for every unit/person in the economy.

In the section above it was mentioned that SRI research agencies analyze stakeholder relations and other extra-financial components to determine a corporation sustainability level. This idea is also accepted in most of the definitions, but it has to be highlighted that until now there is not a uniform definition for Socially Responsible Investments. In alignment with most of the definitions in the market a definition of SRI will follow in the next paragraphs.

Socially Responsible Investments are based on a three dimensional analysis of investments. Sustainable investments do not focus on the short-term financial aspect, but on the ethical and environmental influence of the corporation and the long term financial performance. To

give a good understanding what SRI stands for, we added a definition from well-known academics on this field. Hoepner and McMillan's define SRI as "*investment in capital assets based on screening and selection processes or ownership policies, which are not exclusively developed and practiced on the basis of financial information, but are also developed and practiced on the basis of environmental, social and governance (ESG) criteria that account for the investment's current and future impact on society and natural environment."* (Hoepner and McMillan, 2009, p.18) As earlier mentioned, this is not the only definition in the market, but most of the others are close to Hoepner and McMillan (e.g. Sethi, 2005). Since SRI also focuses on financial performance, it is important that SRI does not get mixed up with Charity Investments. To get a better understanding of the differences we focus on Pearce (2003). According to him, the economy can be split into three sectors as shown in graphic 3. Charitable organizations and social enterprises are part of the so-called "third sector" of the economy. They are not part of the private sector or the public sector. Socially Responsible Investments focus exclusively on the private sector, which is located in the "first sector", and seek for companies that combine economic and sustainable success. So it is clear that charity investments and social responsible investments must be strictly differentiated.

Graphic 3: The three-sector economy system



Source: Pearce (2003), page 25

Different forms can be used to follow a Social Responsible Investments approach. The Ethical Investment Research Service (EIRIS) in cooperation with the UK Social Investment Forum (UKSIF) introduced a paper with the title „The Responsible Investment Approaches of Common Investment Funds“, in which the organization differentiates the three dimensions of SRI. They use the terms support, avoidance, and engagement. At no point of the document they mention that all three pillars have to be followed. So an SRI Fund could only follow one of the three.

»Support« or positive screening focuses on companies with commitment to responsible business practices and positive products/services (EIRIS/UKSIF, 2006, p.5). This includes the best-in-class approach and thematic investment like environmental technologies. Best-in-class approach stands for evaluating all companies in one industry by their extra-financial components, rank them and then to choose only the most sustainable corporations in the industry.

»Avoidance« or negative screening stands for not investing into companies which do not fulfill the social investment criteria (EIRIS/UKSIF, 2006, p.5). If this pillar is followed, certain industries would be excluded out of the universe. This approach varies over different geographical locations. In Austria for example, Nuclear Power is seen as not sustainable, since companies produce nuclear waste. On the other side, Nuclear Power Plants are seen as sustainable in Belgium. They argue that Nuclear Power Plants do not produce any CO2 emissions and are therefore sustainable.

»Engagement« or shareholder activism uses the influence of investors and the right of ownership to encourage more responsible business practice (EIRIS/UKSIF, 2006, p.5). This mainly takes the form of dialogue between investors and companies and may extend to voting practices. These voting practices will have a strong influence on Corporate Governance of a business. The occurring problem from this pillar is that only large financial institutions are able to affect corporations through their large stock holdings. This is why academics state the importance of engagement for institutional investors (Sethi, 2005). Sparkes for example sees the third dimension as obligation in Socially Responsible Investments. In one of his works he states: “SRI is generally considered to be an equity-based activity, as one of its core aims is

to use the power and influence of shareholders to positively affect corporate behavior" (Sparkes, 2001, p. 195).

2.3 The Evolution of Sustainable Investments

Socially Responsible Investments have received enormous attention in the last two decades and today Socially Responsible Investments are present in almost every investment house, but the concept is nothing new. Already several hundred years ago, religious organizations, named the Quakers, used the concept of sustainable investments by using exclusion criteria (Renneboog et al., 2008a). Religious groups still play an important role in the sustainability field. For example Islamic organizations follow their religious tradition (sharia concept) and exclude pornography and gambling out of their investments (Renneboog et al., 2008a). Furthermore, they have special regulations for interest rates. These concepts of sustainability are these days only a small component of the SRI investment idea. The first fund, which used the SRI concept, as we know it today, was launched in 1971 under the name Pax World Fund (Renneboog et al., 2008a). The Pax World fund was established during the Vietnam War and offered investors to exclude corporations with a focus on military and nuclear components. Until the end of the 1990s the SRI approach was strongly value-driven. As Socially Responsible Investments developed further into its modern form, it moved away from an emphasis on ethics and towards incorporating environmental, social and corporate governance factors into investment decisions, thereby becoming an investment strategy that also explicitly seeks investment returns. Investors started to believe that risks could be reduced through an ESG analysis (Deutsche Bank, 2012).

Hudson (2006) outlines four triggers for the rise of sustainable investments. The first one is as mentioned above, religious and ethical believes. The next trigger was mainly established by politicians. It's the social movement, such as the anti-racism movement in the 1960s (Renneboog et al., 2008a). The third component, which accelerated the momentum of sustainable investments, were events with intensive public reaction, as natural disasters and wars. The last but not least component Hudson (2006) mentions is the thoughtfulness from corporations for stakeholders' satisfaction. Furthermore, the society has become more aware of environmental, educational and social difficulties and through the awareness the public desires to have a positive influence on the future. In other words, investors desire a

"feel-good" vibe (Schueth, 2003). The last and most recent driver of sustainability is the momentum it receives through media coverage (Louche and Lydenberg, 2006).

2.4 Main Theoretical Reasonability for Socially Responsible Investments

Research in the SRI field did not increase as quickly as the public interest and capital invested in sustainable investments. Over a long period of time, the majority of academics associated Socially Responsible Investments with underperformance because of a lower diversification possibility. Lately, the interest of academics in that topic increased and the fundamental idea is changing. Though academics still stand on both side and there is no consensus yet.

Renneboog (2008a) claims that there are two main theories, which could explain SRI performance. The first theory mentioned by Renneboog (2008a) which would support underperformance, is the Modern Portfolio Theory. The fundamental idea is to distinguish between two types of risks, systematic and unsystematic risk. Systematic risk is the risk which occurs to the overall market, so it encapsulates the volatility of the fund towards the entire market. This risk cannot be eliminated through diversification and that is why investors are rewarded for this kind of risk (Sharpe, 1964). On the other side, investors do not receive any compensation for taking unsystematic risk, since this risk is firm specific and can be eliminated through diversification (Barnett and Solomon, 2006). Given these risks, academics argue that a portfolio will only lie on the efficient frontier if the unsystematic risk is fully eliminated through diversification. Otherwise, the portfolio will not have the best risk-return profile (Sharpe, 1964). However, if the unsystematic risk is not fully eliminated through diversification, as it is in SRI funds because of exclusion criteria (e.g. military, nuclear power, GMO's, etc.), theoretically they should not lie on the efficient frontier. Hence, these funds should only be able to generate inferior risk-adjusted returns (Renneboog et al., 2008a). Another point academics like to pick on, is the fact that if the Efficient Market Hypothesis holds, SRI Funds shouldn't outperform due to the fact that all ratings and SRI information should be known and absorbed by the market.

The other theory stated by Renneboog et al. (2008a) is contrary to the Modern Market Theory and focuses on asymmetric information. Renneboog et al. (2008a) asserts that SRI managers have a deeper company knowledge whereby they can use this unpublished infor-

mation to outperform the market. Through the greater stock picking ability SRI managers could be able to eliminate companies from their universe, which experience negative financial effects from the likelihood of being involved in costly government regulations, social, and environmental crises (Renneboog et al. 2008a). So he follows the simple idea that analysts and managers are able to find and value environmental, social, and corporate governance risks by analyzing available data. This approach would explain why unsustainable corporations underperform the market, but not why sustainable firms should outperform the market. Further in our study it will be analyzed why the three ESG factors could be performance drivers, but before that, we will test if the ability to analyze firms after ESG criteria produced an out-, under-, or natural performance in the past years.

2.5 Empirical Analysis of Socially Responsible Investments and Performance

By reading through the literature the reader receives all kinds of answers to the question if there is outperformance or not and where this outperformance comes from. Several papers (e.g. Hoepner and Zeume, 2009) argue that corporations with low ESG Ratings must pay higher interests to their investors due to the exclusion of the firm of Socially Responsible Investments (lower demand). However, Haigh and Hazelton (2004) mention the fact that this cannot be correct yet, since assets under management from sustainable investors are not large enough. They give an example and mentions that the largest SRI Fund in the world hardly owns 1% of the shares in each investment (Haigh and Hazelton, 2004). Other criticisms in papers point to the high costs for ESG research and scholars argue that the possible value generated through ESG data is neutralized by the high costs. These examples show how ambivalent the market is by talking about sustainability.

After knowing the theoretical approaches from Renneboog (2008a) and the ambivalency of the market, we will give a short overview on empirical evidence for performance from Socially Responsible Investments.

As mentioned previously the academic world is not consistent yet. Many scholars tested the Socially Responsible Investment funds on their performance. On the one side there are papers that show an underperformance through sustainability. One of the most extensive fund performance studies was done by Renneboog et al. (2008b). In their study they analyzed 440

SRI funds in over 15 countries. The funds were mainly administrated in North America, Europe and Asia-Pacific. They explored risk-adjusted returns through the Carhart multi-factor model and tested for the alpha from SRI funds. As benchmark they used indices and partly conventional fund peers. Furthermore, Renneboog et al. (2008b) adopted Ferson and Schadt's (1996) conditional model to control for time-varying alpha and beta. Overall, the findings suggest that the majority of SRI funds display lower risk-adjusted return than their domestic peers. This would then stand for lower alphas from SRI funds and an underperformance due sustainability. On the other side, you can find studies, which indicate an outperformance. However, most of the studies are not able to find any significance. For example Bauer et al. (2005, 2006, 2007), Cortez et al. (2009), and Galema et al. (2008) use sustainable funds in comparison to different benchmarks and none of them find statistically significant results.

In contrary to the studies above, a study from Barnett and Solomon (2006) focuses on the intensity and style of SRI criteria in sustainable funds. The study aims to test the effect on funds diversification, as well as if superior stock picking can be achieved through certain screening activities. The results state that high SRI restrictions can lead to a lack of diversification, but they also show that certain exclusion criteria avoid investments into low performing industries/stocks. Relating these findings back to theoretical perspective on SRI performance, the authors therefore find empirical evidence and contradiction for the modern portfolio theory.

The problem arising from all above-mentioned studies is the significant human component a fund manager provides to the fund performance. Outperformance/underperformance could be generated through stock picking which does not need to be linked to sustainability of the investments. A better approach to identify the under- or outperformance from ESG criteria would be to analyze the performance on an Index level and that is exactly what Statman (2006) did. His study focused on the American market, where he compared 4 SRI-Indices against the Standard and Poor's 500 Index (S&P 500). His results were all statistically insignificant. In addition to the performance analysis he looked at the composition of SRI funds and found that each fund focuses differently on the single ESG components, especially the Indices use different measurements to identify the social element (Statman, 2006). For example the iShares MSCI KLD 400 Social Index (DSI 400) focuses particularly on the environmental

component and only uses a low number of social elements. Hand in hand, with the findings from Statman (2006), are the findings from Schroeder (2007) who investigated 29 international SRI Funds. So the insignificancy and the factor selection seem to be an international problem.

2.6 ESG Measurements

The ESG selection problem already appeared by comparing different SRI funds and does not end there. Since the industry did not find a common definition it seems to be obvious that this problem also occurs by choosing components to identify the level of sustainability. Only by looking at the largest research agencies in the sustainability field, large differences can be observed. For example MSCI ESG Research analyzes companies ESG risk factors, so they only consider elements, which increase or decrease risks in the industry. In contrast to this, oekom Research selects certain factors for each industry but they also consider not "so relevant" components. Furthermore, Entine (2003) states that such research valuation systems do have fundamental errors. These systems are personally biased, highly selective and do not provide enough information to the outstanding world to understand the reasoning (Entine, 2003). He backs this statement by mentioning the American rating database KLD Inc. (was acquired by MSCI ESG) and notes that there is a lack of explaining the variables, which are used to generate corporate ratings. In addition, Entine (2003) mentions the sources the agencies use to fill their models with data. The main sources for ESG information are CSR Reports from corporations, which lead agencies to transfer text into quantitative data (Entine, 2003). This transfer is associated with individual interpretation of the analyst and thus is exposed to a personal bias. This is not the only criticism the author mentions. He further states that the models, designed by the agencies, are also personally biased since there is no point to refer to. Over all, subjectivity can be observed in the data gathering. On the other side, we have academics, which note that the rating agencies are highly consistent, objective and always follow the same screening criteria (Waddock, 2003). He argues that research agencies always use the same approach and even if they would be biased in their quantitative model, this would affect all screened companies and thus the bias would be irrelevant. Waddock (2003) also criticizes Entines statement; he mentions that Entines assertions are based on no theoretical or empirical evidence.

Finally it can be said that there is no universal framework available yet and it cannot be expected to find a universal framework before an accepted SRI definition is found. Since there is no research on the consistency and objectivity of European ESG Research providers, we assume that the above-mentioned points will also be applicable for the European databases. As long as the results from research providers are not mixed up, we believe that there shouldn't be a problem of a personal bias. To receive a broad overview of the largest ESG research companies and their approach see table 1.

Table 1: Different ESG Rating Agencies

Research Agency	Description
Oekom Research AG[3]	Analyzes all three E, S, and G dimensions with a sustainability focus, whereby a number of variables are identical for all industries and others are industry specific. Inputs are based on publicly available data and questionnaires.
MSCI ESG Inc.	The focus is on risk allocated to each E, S, and G dimension. For every sector/industry key risk factors are identified and analyzed. If a sustainability variable is not a key risk factor, it will not be taken into account.
Sustainalytics	Is a mix between sustainability and allocated risk factors, whereby the sustainability factor outweighs.
EIRIS Foundation	The focus lies on sustainability and is analyzed through 110 different ESG variables. The information is based on publicly available data and dialogues with companies.
ISS A/S	Only provides data for one of the three dimensions. ISS research focuses on corporate governance and therefore part of the agency focuses on active ownership.
CDP Foundation	Collects data on the environmental dimension, with a focus on greenhouse gas emission, water waste and paper & packaging usage.

[3] Further details under 3.2 Data -oekom

3. Empirical Study

After understanding the concept behind ESG and Socially Responsible Investments, we will examine if the insignificancy in SRI fund performance analysis matches a performance analysis on research agency data. In the next section of the paper the focus lies on the chosen data, portfolio formation, methodology, and the results of the head-to-head, Jensen's alpha model (modified CAPM) and Fama and French three-factor model.

3.1 Methodology

The research is split into 3 components. First component is a head-to-head performance approach as research providers usually do in their studies to provide evidence of their research results. A Head-to-Head comparison is a naive approach, which can be simply executed and read. Here, the returns for each constructed portfolio are generated, and then the different portfolios are compared between each other and towards a benchmark. The restriction of a head-to-head approach is simply that it will not provide us with any significant information about the performance, generated through Socially Responsible Investments. Therefore in the second part of the empirical study we will use the advanced CAPM, including the Jensen's alpha to see if our constructed portfolios generate abnormal returns and if they are significant. Furthermore, the advanced CAPM from Jensen was chosen because it's one of the most employed models in the financial research field (Bauer et al., 2007). As mentioned before CAPM stands for Capital Asset Pricing Model and one of the developers was W. Sharpe in the 1960th. The model is based on the modern portfolio theory from Markowitz (1952) and Tobin (1958).The underlying idea, by taking risky assets into account, is that investors are able to reduce their portfolio risk through diversification as long as the risky assets do not reveal a perfect positive correlation. Furthermore, an investor can optimize his portfolio by investing α into a risk free asset and $1 - \alpha$ into a risky asset portfolio, which is diversified. Sharpe took these models, including assumptions and further assumed that all investment opportunities can be traded in one market, all investors can trade on the financial market, and all investors have the same full information. These assumptions lead to the result that there will only be one efficient portfolio for all investors, here called market portfolio. This is due to the fact that all market participants have the same information, and no investor would hold overpriced assets respectively sell underpriced assets. So overall the

market would be perfectly priced (all assets lie on the security market line) and holding all assets in the market would minimize the risk, therefore every investor would hold the market portfolio as its risky asset and manage their risk aversion through the combination of the market portfolio and a risk free asset. So the result from Sharpe was that the risk, which comes from a single risky asset (unsystematic risk) can be eliminated through the diversification and an investor will only need to hold the market risk (systematic risk). Therefore, investors should only be rewarded for the systematic and not the unsystematic risk. Since all assets are represented in the market portfolio an assets' sensitivity towards the systematic risk can be measured through the Capital Asset Pricing Model. Over all the CAPM shows how sensitive an asset is towards the systematic risk in the market. To receive a better understanding the equation of the model follows in Equation 1:

$$E(r_i) = r_f + \beta_i\big(E(r_m) - r_f\big) \qquad [1]$$

where $E(r_i)$ stands for the expected return of an risky asset, $r_f$stands for the risk free asset, $E(r_m)$ for the expected market return, and β_i represents the systematic risk allocated to the risky asset.

Jensen (1968) was interested in the question how to measure the ability of portfolio managers and he believed a fund manager could outperform the market. To find an answer to his question he added to the basic Capital Asset Pricing Model the stock picking abilities of the manager. Since Jensen focused on the valuation of managers and their stock picking in their portfolios, he replaced the expected return of the market portfolio through the benchmark of the portfolio manager. The final equation from Jensen can be seen in Equation 2.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i\big(r_{m,t} - r_{f,t}\big) + \varepsilon_{i,t} \qquad [2]$$

where $r_{i,t}$ is the return of a risky asset (risky portfolio) at time t, $r_{f,t}$ stands for the risk free rate at time t, the benchmark is seen as market return at time t and is represented through $r_{m,t}$, and the Jensen's alpha (1968) is represented through α_i. The beta coefficient[4] (β_i) captures the asset (portfolio) systematic risk exposure to the benchmark, and ε_t, the error term, specifies idiosyncratic returns, which cannot be explained by the market proxy or the specific portfolio return (Hoepner and Zeume, 2009). Standard errors are designed to control for

[4] $\beta_m = \frac{\sigma_{i,M}}{\sigma_M^2}$

autocorrelation and heteroskedasticity. In our approach we rather focus on portfolio return than on risky assets, we use the 3-month EURIBOR as risk free rate, and our benchmark is the MSCI World Index. So we will use the modified CAPM by Jensen and focus on the Jensen's alpha to see if the market explains the total portfolio performance or if sustainability can lead to an outperformance of the market, which would also be an argument against the efficient market hypothesis.

After receiving any significant results the question appears, if the out- or underperformance was generated through an SRI approach, or if other stock characteristics are responsible for the performance. To identify the relevant performance drivers the Fama and French model (1993) is used. Fama and French went one step further then Jensen and included two more factors into the basic CAPM model then Jensen did. In their research they found out that small companies and companies with a lower book to market ratio, also called value stocks, outperform the market. Thus, they added these factors into Jensen's Model and created the three-factor-model. In their study they mentioned that Jensen's model only explains 70 % of the data und that the three-factor-model explains more than 90% of the data (Fama and French, 1993). Several other academics also tried to develop similar models to explain outperformance through different firm characteristics, but one of the most recognized ones is still the three-factor-model. The recognition is based on the significance of the three factors, the good fit on data and that the small firm and value firm factor works in a large number of markets. Due to these reasons the Fama and French model should provide us with more information concerning the performance and is a good addition to the modified CAPM model from Jensen. To understand the differences between these two models the basic three-factor model from Fama and French (1993) will be explained below, whereby the model was directly adjusted by us for portfolio positions.

This model is estimated by

$$r_{p,t} = \alpha_p + \beta_m\left(r_{m,t} - r_{f,t}\right) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) + \varepsilon_t \qquad [3]$$

where the excess return $\mathrm{r_{p,t}}$, return of the market $\mathrm{r_{m,t}}$, the Jensen's alpha α_p, the beta β_m, and the error term are identical to the factors in the Jensen model. The β_{SMB} identifies how much performance can be allocated to the outperformance from small firms in the portfolio. β_{HML} is similar, just that it accounts for the outperformance generated through value firms

in the portfolio. How the style effects (HML and SMB) were constructed will be explained below.

To be able to conduct the different models further factors have to be calculated. To compute the excess return for each portfolio in the Jensen model, we downloaded weekly total returns from Bloomberg from January 2006 until December 2013. To finally compute the excess return for each portfolio we subtracted the risk free rate from each portfolio return. Our risk free rate is the three-month Euribor, so we downloaded the 3M EURIBOR from Bloomberg and transferred the 3M EURIBOR with the following formula:

$$r_{f,week} = (1 + r_{3M})^{\frac{1}{52}} \quad [4]$$

into a weekly rate. The same weekly risk free rate is used to generate the market premium by subtracting the risk free rate from the weekly MSCI World Index return.

For the Fama and French model, in addition to the market premium and the risk-free rate, which are calculated as in the modified CAPM from Jensen, we computed the styling factors SMB and HML. For the SMB factor we used the MSCI World Small Cap Index and calculated the abnormal return against the standard MSCI World Index. For the HML factor we followed the same procedure, but used different Indices. Here we downloaded the MSCI World Value Index and MSCI World Growth Index from Bloomberg and generated the weekly abnormal return. All data are downloaded in weekly terms and with dividends.

For all portfolios and approaches, the chosen benchmark is the MSCI World Index, since the companies analyzed by oekom are all listed and spread over the world with a focus on Europe and the US. More than 75% of all companies in the MSCI World Index are covered by oekom. The benchmark was continually (monthly) adjusted to avoid any mispricing over time.

In most of the past SRI performance studies, as mentioned above, funds with a special focus in one of the three ESG components were analyzed. To work around this problem the focus of our research lies on data from a selected research agency, which provide ESG research for an international universe and with no tendency in a specific field. Furthermore the use of research agency data will also eliminate any bias provided from managers stock picking ability.

3.2 Data

Prior SRI research had two major shortcomings. One of the largest difficulties by analyzing companies, based on ESG criteria, was the availability of data. Since the topic ESG only became famous at the end of the 1990s, large ESG research companies were founded around those years, but the development of a proper quantitative model took about a decade and is today still not unique. Furthermore, some consolidations have taken place in the industry and quantitative models where merged and adapted. The other shortcoming of prior SRI research is the market focus. Most studies focused on the US Market and have also used US based research agencies. We contacted numerous research providers with the demand to contribute data over a longer time horizon with an international market focus. Most providers only had data from the mid- or end-00th and had a market focus at its inception. The best suitable provider was oekom, since the company provides corporate ESG ratings for an international stock universe, has frequent data for the last 9 years available, reasonable prices, and the methodology behind the ESG criteria, used to identify sustainable investments, has a fundamentally strong base.

oekom research AG is a public corporation with approximately 50 independent private stockholders. According to the CEO of oekom research AG this distribution of shareholders ensures the independency of the company (oekom, 2014a). As mentioned above, oekom provides company ratings on a quarterly base from end 2005 onwards. Nowadays, the research provider covers approximately 3000 companies worldwide[5]. The coverage is based on three different pillars: important national and international stock indices (e.g. MSCI World Index), companies in sectors with a high sustainability reference, and other non-publicly traded firms (mainly used for the bond market).

oekom's rating methodology is separated into a social and an environmental component. The social component also contains the corporate governance dimension. Since industries are exposed to different sustainability factors, every industry has in average 100 industry-specific indicators to identify the sustainability of the firm. To ensure a high degree of accuracy, all indicators follow a pre-defined concept (Frankfurter-Hohenheimer Leitfaden), are weighted, tested, and aggregated to generate a total score for each company. Defined indicators are then reviewed over a 1 year basis. The score is then transferred for the consumer

[5] Data from February 2014

into a performance grade system from A+ to D-. The best SRI rating a firm can achieve is an A+ and the worst is a D-. oekom uses a variety of inputs to form their score. The first step is to analyze public available information. After screening CSR reports, websites, and annual reports the first rating draft is send to the analyzed company to give them the opportunity for comments. In the next step the comments are examined and if they fall into weight they are accounted. Additional to the above mentioned procedure, oekom cooperates with a number of NGO's to receive substantial information of firms. These cooperations are mainly used to examine exclusion criteria (e.g. controversial weapons), incidents of corruption, etc. (oekom, 2014a). Furthermore, the research agency detects the best companies, determined through their ESG ratings, to generate a Best-in-Class ranking for their clients. All companies which belong to the Best-in-Class are so called Prime-Companies. oekom already produced a head-to-head study which showed the outperformance of the Prime-Companies (oekom, 2014b).

3.3 Portfolio Formation

In order to provide a good sustainable research, we decided to analyze every performance grade (A - D) from oekom. In the next paragraph, this process will be explained in detail.

Data, provided by oekom, include ratings on quarterly basis of a range of companies in the covered universe from 2005-2014. During the sample period the universe expanded from 600 to 3000 companies. First we had to clean the data in terms of availability on Bloomberg. We mainly cleaned out data with missing ISINs or companies covered by oekom, which are not publicly traded. For traded companies we generated a time series including stock price, book-to-market ratio, and market capitalization for the entire time horizon. Afterwards, we structured a portfolio for each performance grade. To ensure a diversified portfolio with a sufficient amount of companies we decided to cumulate portfolios of one grade. In other words, we pooled all "plus", "neutral", and "minus" rating of one grade (e.g. B+, B, and B-) into one total cumulative rating (e.g. B). Furthermore, we spared to build a portfolio A, for the sake of the small number of A-rated (incl. A+ and A-) companies as you can backtrace in Table 2. The number of companies with an A rating fluctuated between 9 and 22. Companies with a B rating therefore represent the most sustainable portfolio, whereas portfolio D represents the least sustainable and C ratings are in between. The portfolios are updated at the

beginning of each quarter following the rating information of oekom research AG. So the data at the end of Q4 2005 were used to construct the portfolio in Q1 2006. For each of the quarterly updated portfolios (B, C, and D portfolio) we observed the performance through the stock price development. Weights of the companies are set up according to their relative market capitalization[6] in the according portfolio, at the beginning of each quarter. So we generated value weighted portfolios. Thus, depending on their weight in the portfolio, the stock price development of each and every company contributes to the performance of a portfolio in a quarter. Afterwards, the portfolio gets restructured if ratings have changed. For a better understanding how the number of companies in every portfolio changes, we provide Table 2 as an overview. Due to small changes over the year we only included the data at the beginning of each year and no quarterly data.

Table 2: Rating Development of the oekom Universe

Date	01.2006	01.2007	01.2008	01.2009	01.2010	01.2011	01.2012	01.2013	12.2013
# A-rated	22	17	9	10	14	17	18	21	15
# B-rated	84	105	146	162	181	205	203	227	234
# C-rated	250	280	327	386	419	459	499	514	590
# D-rated	42	45	49	42	47	40	54	73	715

Once the constituents of all portfolios have been completed, the performance is measured by linking the previous downloaded Bloomberg data, stock prices, and market capitalization to the companies in the portfolios. To provide a sufficient detailed performance analysis we considered equity prices with dividends and without. To ensure the completeness and the comparability of portfolios we finally used stock price performance including dividends because they display the total return of a stock and do not only take price changes into account. To be consistent in our research and to be able to compare the portfolios, we converted all equity prices into our core currency US dollars. We chose the US dollar because of the high number of stocks denominated in the currency. We are aware that our samples

6 $\frac{P_{i,t} * number\ of\ Stocks_{i,t}}{\sum_{i=1}^{n} P_{i,t} * number\ of\ Stocks_{i,t}}$

may suffer from a survivorship bias, since we adjusted our portfolios for companies, which merged or disappeared from the equity market. At this point, our argumentation is in line with Kreander et al. (2005) who states that the survivorship bias would affects all portfolios in our sample and therefore, the overall influence on our results should be minor (Brown and Goetzmann, 1995). Subsequently, we calculated the total return over the time horizon for each firm and we computed the return for the value-weighted portfolios. For this we rebalanced the portfolio in each quarter concerning their sustainability rating and furthermore, at the same time the value-weighted portfolios are rebalanced due to the relative market capitalization, which we explained earlier.

For a better understanding of the usability of sustainable rating, the sample is split into two subsamples. Subsamples can provide the information, in which kind of market a sustainable investment approach is more valuable. To identify two different types of markets, we split the whole sample at the beginning of March 2009. This provides us with a pre-crisis and an after-crisis sample, so we can analyze sustainable ratings in a bull and a bear market. Both samples are constructed out of the total period sample. Thus, they don't need further explanation. The pre-crisis sample begins at the 2nd January 2006 and ends with the first March 2009. The after-crisis sample starts in March 2009 and ends in December 2013.

Table 3 reports descriptive data of the constructed portfolios. This descriptive table contains, among others, the returns, standard deviation of returns , Sharpe ratio of the different portfolios, and the three different samples. To calculate the returns the continuous approach, with the following formula, was used:

$$r_{p,t} = \sum_{i=1}^{n} w_i * LN\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \; ; \; \sum_{i=1}^{n} w_i = 1 \qquad [5]$$

Furthermore, the Sharpe Ratio can be seen as a first indicator for the risk-adjusted performance. In essence, it shows the risk-return relationship of a portfolio, identifying how much extra risk an investor has to take to receive an additional unit of return. It measures the average excess return over the risk-free rate divided by the standard deviation of the portfolio. The outcome can be seen as the higher the better.

$$SP_p = \frac{r_{p,t} - r_{f,t}}{\sigma_p} \qquad [6]$$

Table 3: Descriptive Statistics

Portfolio	Total Return	Average Year to Year Return	Standard Deviation Year to Year	Maximum Return	Minimum Return	Sharpe Ratio*	Skewness	Kurtosis
Total period sample								
B Rating	0.4013	0.0502	0.2214	0.1112	-0.2347	0.1305	-1.3163	11.1704
C Rating	0.1388	0.0174	0.2076	0.1190	-0.2387	-0.0189	-1.5372	13.6870
D Rating	0.1551	0.0194	0.1967	0.1083	-0.1957	-0.0095	-0.8975	9.0089
Pre-crisis sample								
B Rating	-0.2724	-0.0908	0.2743	0.1112	-0.2347	-0.3324	-1.5229	11.4763
C Rating	-0.4603	-0.1534	0.2566	0.1190	-0.2387	-0.5994	-1.8179	14.3917
D Rating	-0.3214	-0.1071	0.2278	0.1083	-0.1957	-0.4720	-1.2904	11.1800
After-crisis sample								
B Rating	0.6280	0.1256	0.1917	0.0863	-0.1018	0.6116	-0.4612	4.6228
C Rating	0.5582	0.1116	0.1797	0.0836	-0.0982	0.5745	-0.4491	4.7669
D Rating	0.4365	0.0873	0.1847	0.0975	-0.0822	0.4273	-0.2327	4.2779

*Average yearly Sharpe Ratio

3.4 Results

3.4.1 Head-to-Head

Table 4 exhibits the average weekly return, average yearly return, relative performance over the whole sample, and the relative yearly performance for each portfolio in the different samples. The portfolios are value weighted and portfolio B is, as mentioned before, constructed out of companies with the best ESG ratings. Portfolio D consists of the worst ESG firms and portfolio C is in-between. In addition, the standard deviation of yearly returns is included in the table to rate movements and the risk allocated to each portfolio. The given data shows an outperformance of the B portfolio against the other ESG portfolios in all samples. In two out of the three samples, the B portfolio also beats the benchmark. In the total period sample, we observed an outperformance of 0.1823 over the full period, which can be broken down, by taking the reinvestment into account, to an average yearly outperformance of 0.0212. Portfolio C and D both perform similar poorly and lie with -0.0801 and -0.0638 behind the benchmark and therefore, underperform portfolio B by 0.2624 and 0.2461. In the subsamples we receive a different picture, concerning the benchmark and portfolios. In the pre-crisis sample, which can be linked to a bear market, all portfolios beat the benchmark, whereby portfolio B shows a relative yearly performance of 0.0547 and portfolio C is the worst performer in this sample with a relative yearly performance of 0.0003. It has to be mentioned, that Portfolio D performance better then portfolio C with a relative yearly performance of 0.0411 compared to the benchmark, which can be observed in Table 4, column 5. Interesting is the volatility and movements of the portfolios in the full and pre sample. The two portfolios with low ESG Ratings move with the benchmark and portfolio B has higher volatility in its chart (Appendix IV.I and IV.II), whereby the number of companies in the portfolios is stable and large enough to ensure a diversified portfolio. In the last sample, the after-crisis sample, none of the portfolios is able to outperform the benchmark, though portfolio B is close to it with an underperformance of -0.0105 and this time, the portfolio is less volatile and moves similar to the market. Portfolio D is a clear underperformer in a bullish market with a relative performance of -0.0488 and therefore, underperformance portfolio B and portfolio C. Portfolio C lies with a relative performance of -0.0245 between portfolio B (relative performance of -0.0105) and D. So it seems that in the head-on-head approach, sustainable investments lag in a bullish equity market and outperform in bearish markets. Furthermore, portfolio B always outperforms the other constructed portfolios. Portfolio C

and D are not in line with our idea that a higher ESG Rating brings a better performance, since portfolio C underperformed in two out of three samples.

The average number of companies shows a higher number in the C portfolio and through a more detailed analysis; we observed that more industries are represented in the C portfolio. This result could show us the inconvenience of the lower diversification argument, which is often brought up by academics or ESG detractors. Furthermore, the average number of companies included in particular portfolios demonstrates the fortitude of the analysis. So overall, the head-to-head record shows an outperformance through high ESG ratings and is therefore in line with the results from Renneboog et al. (2008b), but this is only a direct comparison and wasn't yet tested on its significance.

Table 4: Results Head-to-Head Analysis

Portfolio	Average Weekly Return	Average Yearly Return	Standard Deviation Year to Year	Relativ Performance (yearly)	Relativ Performance (overall)
Total period sample					
B Rating	0.0009	0.0502	0.2214	0.0212	0.1823
C Rating	0.0003	0.0174	0.2076	-0.0104	-0.0801
D Rating	0.0004	0.0194	0.1967	-0.0082	-0.0638
Benchmark	0.0005	0.0274	0.1948	-	-
Pre-crisis sample					
B Rating	-0.0017	-0.0908	0.2743	0,0547	0.1889
C Rating	-0.0030	-0.1534	0.2566	0,0003	0.0011
D Rating	-0.0021	-0.1071	0.2278	0,0411	0.1399
Benchmark	-0.0030	-0.1538	0.2399	-	-
After-crisis sample					
B Rating	0.0024	0.1256	0.1917	-0.0022	-0.0105
C Rating	0.0021	0.1116	0.1797	-0.0052	-0.0245
D Rating	0.0016	0.0873	0.1847	-0.0105	-0.0488
Benchmark	0.0026	0.1361	0.1634	-	-

3.4.2 Jensen's alpha

The significance of the portfolios under- and outperformance is initially tested on the basis of a t-statistic with the Jensen's alpha. The full results can be found in Table 5. The results are mixed, but overall, our results are mainly consistent with findings from Cortez, Silva, and Areal (2009) and Galema, Lensink, and Spierdijk (2008), where no clear significance was identified. The R^2 of all samples and portfolios are quite good and explain most of the data. They lie between 0.85 and 0.98. Furthermore, the performance can mainly be explained by the market return. The market coefficient lies close to one in all observations.

Results from the total period sample are partly as we expected. Portfolio B has the highest positive alpha (0.000386), but portfolio C has the lowest negative alpha with a value of 0.000238 and not portfolio D (-0.000178). So by following our idea - buying companies with the highest rating and selling them with the lowest rating – we could generate an weekly alpha of 0.000564, which would lead to an yearly alpha of 0.0293. But it has to be mentioned that a higher alpha could be reached by going short on portfolio C. As nicely the results sound, all generated alphas do not show any significance in the full period sample. The results for the pre-crisis sample are similar to the one of the total period sample. Portfolio B generates the highest alpha with a value of 0.001361, which would lead to an outperformance of 0.07077 over the year through the alpha. Portfolio C and portfolio D have both lower alpha values, but both are positive as it can be seen in Table 5. Portfolio C has a value of 0.000089 and portfolio D generates similar to the total period sample a higher alpha then the C portfolio (0.000354). So again, we have the same order as in the total period sample. The difference here is that we found a significant alpha for our B portfolio in the pre-crisis sample with a t-statistic of 1.7909 and a p-value of 0.0751. In contrast to the total and pre-crisis sample, the after-crisis sample results is in line with our beliefs. Here, portfolio B outperformances C and portfolio C outperformances D, whereby all alpha values are negative and underperform the benchmark. Again, we find one portfolio with a significant alpha. This time, portfolio C has a significant alpha of -0.000558 and a t-statistic of -2.4770.

Table 5: Results Jensen's alpha Analysis

Portfolio	alpha	t-Statistic	Market Beta	t-Statistic	R^2
Total Period Sample					
B Rating	0.000386	0.9758	1.0800***	78.0019	0.9363
C Rating	-0.000238	-1.1713	1.0366***	145.7223	0.9808
D Rating	-0.000178	-0.3387	0.9184***	49.8836	0.8574
Pre-crisis Sample					
B Rating	0.001361*	1.7909	1.0680***	48.6769	0.9345
C Rating	0.000089	0.2352	1.0239***	94.0704	0.9816
D Rating	0.000354	0.4236	0.8631***	35.7446	0.8850
After-crisis Sample					
B Rating	-0.000393	-0.9364	1.1037***	61.8220	0.9398
C Rating	-0.000558**	-2.4770	1.0572***	110.2501	0.9802
D Rating	-0.000916	-1.3778	1.0042***	35.4522	0.8369

P-values: *p < 0.10; **p < 0.05; ***p < 0.01

So overall, we only found significant results in the pre- and after-crisis sample for two different portfolios. After the Jensen's alpha test, it can be said that no significant underperformance or robust outperformance is apparent since we only found a significant outperformance form portfolio B in the pre-crisis sample and a significant underperformance of portfolio C in the after-crisis sample. Thus, we shouldn't reject the hypothesis from scholars that well ESG-rated firms have fewer risks in a crisis than firms with bad ratings, since all portfolios outperformed in the pre crisis sample and one of them was significant.

3.4.3 Fama and French

To identify what the performance drivers are, we split the Jensen's alpha, by applying the Fama and French model (1993), into different components. In the total period sample, the value of alpha decreased in portfolio B and increased in portfolio C and D, whereby the alpha values from portfolio C and D remain negative, as it can be seen in Table 6. However, the order is identical to the Jensen's alpha results; Portfolio B outperforms the others and portfolio D and C are close to each other. Now, Portfolio B has a weekly alpha of 0.000274 (0.014248 p.a.), which cannot be declared as significant with a t-statistic of 0.7474. The "small minus big" and "high minus low" coefficients are both negative and significant in this equation. The SMB coefficient has a value of -0.2591 and it can be said that our portfolio B includes more large-capitalized firms than small-capitalized firms, which could stand for the fact that large firms have better ESG ratings or small firms do not receive any rating and therefore are not included in the universe. The HML factor has a value of -0.3978 and indicates an excess of growth firms in the portfolio. Overall the model explains 94.58% of the data (R^2 of 0.9458). The same accounts for portfolio C and D were we found an R^2 with a value of 0.9803 and 0.8585, but the significance of the coefficients in the model for C and D are mixed. Portfolio C is similar to portfolio B, were all coefficients, beside the alpha, are highly significant and the coefficient for SMB is negative (-0.1401) and the HML coefficient is positive (0.1129). In portfolio D, only the market coefficient is significant with a value of 0.9145.

In the two subsamples, the results are similar. The pre-crisis sample also explains more than 88% (R^2 between 0.8876 and 0.9848) of the model and shows positive alpha values for all portfolios. In contrary to the results of the Jensen's alpha model, the alpha of portfolio B is

no longer significant with a value of 0.000920 and a t-statistic of 1.4181. The algebraic signs and the significance of the control variables are identical to the total sample period and can be seen in the Table 6. So as mentioned before, some academics (e.g. Pinner, 2007) believe that sustainable investments have lower down side risks and this belief finds some support in the results of the pre-crisis sample, whereby we only have a significant portfolio by using Jensen's model.

In the after-crisis sample we again find high R^2 for all portfolios, whereby the model from portfolio D only explains 84.50% of the data. The alpha of portfolio B outperforms the value of C and the C portfolio outperforms the one of D, which is in line with our hypothesis and our previous after-crisis results. Similar to our Jensen model approach, the alpha from portfolio C is, with -0.000373 and a t-statistic of -1.8024, significant. The other portfolios didn't show a significant alpha, as it can be seen in Table 6. By looking at the market, SMB, and HML factors we see that all coefficients, beside the SMB from portfolio D, are significant and show that all portfolios contain more large-capitalized firms then small-capitalized firms. Furthermore, the HML coefficients are mixed. Portfolio C and D have positive HML coefficients, which signals more value than growth firms and portfolio B is via versa.

To summarize the findings, we were not able to find a unique result. It can be said that portfolios, constructed through ESG ratings do not generate a negative or positive performance because the significance of portfolio B and portfolio C, in two different modified CAPM samples are not sufficient to define an outperformance for the whole ESG approach. However, a new hypothesis can be stated through our results. After finding only negative coefficients for the small-minus-big variable, it seems to be more likely that large firms receive a sustainable rating from research agencies because they report more. Furthermore, companies with a high ESG rating could have less risk in bear markets.

Table 6: Results Fama and French Analysis

Portfolio	alpha	t-Statistic	Market Beta	t-Statistic	SMB Factor	t-Statistic	HML Factor	t-Statistic	R^2
Total period sample									
B Rating	0.000274	0.7474	1.1269***	80.4004	-0.2591***	-4.9345	-0.3978***	-8.2478	0.9458
C Rating	-0.000126	-0.6728	1.0347***	144.5176	-0.1401***	-5.2218	0.1129***	4.5823	0.9803
D Rating	-0.000112	-0.2134	0.9145***	45.4528	-0.0574	-0.7610	0.0841	1.2150	0.8585
Pre-crisis sample									
B Rating	0,000920	1.4181	1.1275***	54.2482	-0,3662***	-4.4908	-0,5799***	-8.0418	0,9532
C Rating	0,000048	0,1387	1.0158***	95.1983	-0,1554***	-3.5793	0,1027***	2.6678	0,9848
D Rating	0,000200	0,2391	0,8762***	33.9391	-0,1972*	-1.8778	-0,1144	-1.2318	0,8876
After-crisis sample									
B Rating	-0.000347	-0.8393	1.1369***	57.3704	-0.2060***	-2.9875	-0.1789***	-2.7970	0.9428
C Rating	-0.000373*	-1.8024	1.0675***	107.6454	-0.1647***	-4.7722	0.1195***	3.7343	0.9837
D Rating	-0.000701	-1.0681	0.9875***	31.3989	-0.0376	-0.3435	0.3339***	3.2889	0.8450

P-values: *p < 0.10; **p < 0.05; ***p < 0.01

4. Possible Performance Drivers from ESG Ratings

Almost all studies mentioned in our work, and also our empirical study, were based on one rating, which was constructed out of different ratings in the environmental, social, and corporate governance area. Thus it cannot be said if one of the ESG dimensions provides outperformance and another one underperformance. To bring light into the darkness we tried to define potential performance drivers and inhibitors for each dimension. This will be done through a literature review and the corresponding empirical studies, where we define how E, S or G can be measured and if it brings performance. We will work through every dimension and present different approaches. It has to be mentioned that the E, S, and G dimension is differently covered through literature, so we tried to select the most relevant ones. We will start with environment, followed by the social dimension, which is sparsely covered, and finish with the most covered dimension corporate governance.

4.1 Environmental component in ESG criteria

Environment is a term everybody hears day by day in the news. We learn about environmental disasters, which reach from hurricanes up to chemical spill outs. The latest incident was in Turkey (2014) where a mine collapsed over a number of mineworkers. Another worldwide known catastrophe in which a company was involved was the nuclear disaster in Fukushima. A Tsunami hit the Japanese coast and destroyed a nuclear power plant. Since the plant was not fully up-to-date concerning technical standards, Tesco, the provider of the plant, had to release nuclear-contaminated water into the ocean. This sample shows that the whole society is affected by the behavior of firms concerning the environment, but do investors value firms' efforts in environmental projects to prevent such events?

In 1994 US based firms used more than $120 billion to obey with environmental laws (Vogan, 1996). This amount is approximately 2% of gross domestic product (GDP) in 1994, what seems to be a high effort to protect the environment (Konar and Cohen, 2001). Some firms go even further and reduce their environmental impact beyond the law, but do these expenses generate value to shareholders? In the past, academics assumed that environmental protection from firms, provides none to just a few financial benefits to firms and inves-

tors, but this believe has changed dramatically in the last two decades (King and Lenox, 2001). Nowadays, many scholars claim that there is a relation between improvement in environmental performance and financial benefits (Nelson, 1994; Hart, 1997; Konar and Cohen, 2001). They state that environmental investments are a win-win situation for both, the firms and the environment, which means that shareholders have a double benefit. Jacobs, Singhal, and Subramanian (2010) split the benefits from environmental performance into revenue gains and cost reduction. They mention that through a higher environmental performance than the law requires, companies can increase profits in existing markets and/or new markets. Higher profits in existing markets can be realized through an increase in the firms and the products reputation (Klassen and McLaughlin, 1996). Furthermore, new markets can be approached through the increasing demand from customers for more eco-friendly or energy saving products (Porter and van der Linde, 1995). For example the US federal agencies are required to implement environmental criteria in their product selection (EPA, 2008), so firms with a higher eco-efficiency could be new providers to the US federal agencies. On the cost side it is straightforward that environmental performance can reduce the amount of waste and the cost from material input usage, including energy and water (Sroufe, 2003). Whereby, the reduction of components would furthermore bring lower transportation costs. A side effect of reducing environmental burden is the innovation of firms. Through investing in environmental performance, the company stays up-to-date and can generate a competitive advantage (Russo and Fouts, 1997). With a higher environmental performance, a firm can also reduce its reputation risks, when crisis or disasters appear and costs for lawsuits should be decreased though higher eco-efficiency (Karpoff et al., 2005). Dowell et al. (2000) add that development and maintenance cost are lower through continuous environmental updates and firms' innovation can increase employee's morale and productivity, but it has to be hold in mind that environmental investments can be risky and costly if they are implemented too early (Dowell et al., 2000). All the above-mentioned facts can be directly allocated to higher environmental performance. Besides them, there could also be an indirect effect of higher environmental performance. Financial analysts could be willing to recommend the firm because of their environmental performance and investors could be willing to hold their holdings for a longer time because of the competitive advantage and improving cost revenue spread (Konar and Cohen, 2001).

Theory seems to provide us with a row of arguments why environmental performance should lead to stock outperformance of firms, but apparently many academics complain about the usability of the research material and their data. Konar and Cohen (2001) separate the research on environmental performance into two segments: comparing financial to environmental performance over time and analyzing the environmental performance to their firms' market value mainly through event studies. As first approach they state that past studies had difficulties with objective environmental factors and up-to-date data, since most of them used the industry studies from the Council on Economic Priorities (CEP), which examined pollution records for different industries in 1970. A number of studies found out, that the control of pollution had a significant positive effect on financial performance of firms (e.g. Spicer,1978), but as mentioned, these studies were based on old data. Furthermore, they criticize the small samples used in the second approach and the usability of these event studies. For example, Klassen and McLaughlin (1996) analyzed the financial performance of firms connected to environmental news and receiving environmental awards. This study seems to be redundant, since positive news will most likely increase stock prices and negative news will lead to a stock decline, so it doesn't matter on what type of news it is based. So the question arises, how the environmental effort of a firm can first of all be measured and secondly, how it can be linked to stock performance or enterprise value.

4.1.1 Usage of Tobins Q to identify if the market values environmental performance

As mentioned above, in the past it was hard to receive qualitative data to measure environmental performance. In the last decade, changes have taken place in the provision of environmental data. Many industries are expected to report their CO2-emission from their production and/or end-goods. For example, in Europe one of the most affected industries is the Automobile industry, were in 2021, cars will not be allowed to produce CO_2 emission of more than 95 grams per kilometer.

In the US, one of the first steps to provide more data about environmental performance was made in 1987/88, where agencies demanded the market to report about their firms toxic release inventory (TRI). All facilities must complete annual TRI Reports if they manufacture or process 25,000 pounds, use more than 10,000 pounds of any listed chemical during a cal-

endar year, and employ at least 10 full-time employees (King and Lenox, 2001). As soon as this information was available, scholars started to use and analyze them. Hamilton (1995) for example analyzed 436 companies in the US, which had to report their TRIs and found significant negative abnormal returns. Furthermore, Konar and Cohen (1997) selected companies with the strongest stock increase/decline on the TRI announcement and looked at the TRI improvement. They found out that firms with the strongest abnormal return reduced their TRI emission more than the average in the industry. To see how intensive the environmental performance affects firm value, Konar and Cohen (2001) wrote a study where they implemented the toxic chemical releases into a Tobin's Q calculation.

The financial world defines Tobin's Q as the ratio between a firm's market value and firms' replacement costs for their tangible assets. A firm's present market value can be disaggregated into intangible and tangible assets as it can be seen in equation 6. So if a company would have zero intangible assets the Tobin's Q should be 1 as seen in equation 7.

$$MV = V_T + V_I \quad [7]$$

$$q = Tobins\ Q = \frac{MV}{V_T} = 1 + \frac{V_I}{V_T} \quad [8]$$

Konar and Cohen (2001) use the Tobin's Q to identify how much influence environmental factors have on the firm value. For Konar and Cohen, environmental information is defined as intangible assets, so the focus of their study concentrates on them. Since there are more components then TRI that affect intangible value, they took them into account, which will be explained later on. Based on their literature review, Konar and Cohen determined the following factors which affect market value: market share of firm, sales growth, research and development intensity, import-consumption ratio, and industry concentration ratio. Furthermore, they added toxic chemical emission and the amount of environmental lawsuits pending on the firm in 1989 for the environmental factor (Konar and Cohen, 2001).

After identifying their independent variables for the regression, they sourced data from all S&P 500 companies, whereby they cleaned out all companies, which are not affected by any

toxic chemical emission (banks and insurances). After eliminating the irrelevant industries and companies with missing data, they had a sample of 233 companies.

Concerning the independent control variables, the results found by Konar and Cohen (2001) are generally as expected and are similar to the results in other known sources (Konar and Cohen, 2001). By looking at the usual value drivers, their Tobin's Q is positively related to research and development, growth in revenue, market share, and advertising expenditure. Whereby they mention that the adverting variable could be related to environmental performance, but do not reveal any detail. After analyzing the usual performance drivers, they focused on the effect of environmental components on firms' intangible asset value. By looking at the independent variables, toxic chemical emission (TRI) and the number of lawsuits, in the regression it can be said that their influence is negative on firms Tobin's Q. These results are highly significant as it can be observed in Table 7. Konar and Cohen (2001) then do a number of robustness tests. All of them show the same significant results for the two environmental variables and their magnitude is relatively stable. Furthermore, the scholars check the influence on the different industries since different industries have different levels of toxic emissions, but over all it can be said that a 10% decrease in TRI emission would increase the intangible asset value by $34 million (Konar and Cohen, 2001).

The study from Konar and Cohen (2001) gives the feeling that environmental factors could influence market value of a firm, but the sample they used is actually based on one year and misses the long-term effect. The TRI data only came out in 1988 and shareholders could have been shocked, so that the market reaction could have been only a value correction. Compared to Konar and Cohen (2001), King and Lenox (2001) looked at the years between 1987 and 1996 and used a similar approach.

Table 7: Effect of Environmental Performance on Intangible Firm Value from Konar and Cohen (2001)

Dependent Variable	(1) (q - 1)	(2) ln (q)	(3) (q - 1) (interaction terms included)	(4) ln (q) (interaction terms included)
Constant	1.83	0.84	2.40	1.04
	(0.001)***	(0.000)***	(0.000)***	(0.000)***
LN of replacement cost of tangible assets (LNRV89)	-0.20	-0.08	-0.25	-0.10
	(0.003)***	(0.001)***	(0.000)***	(0.000)***
Advertising expenditures as percentage of RV89 (ADVAL89)	0.005	0.002	0.03	0.01
	(0.12)	(0.08)*	(0.002)***	(0.003)***
Advertising expenditures * growth in revenue (AD87789)			0.002	-0.002
			(0.91)	(0.79)
R&D expenditure as percentage of RV89 (RDVAL89)	8.08	1.98	5.08	2.40
	(0.001)***	(0.002)***	(0.24)	(0.22)
R&D*growth in revenue (RD8789)			8.08	4.09
			(0.009)***	(0.005)***
Market share (MSH89)	1.42	0.69	1.23	0.63
	(0.002)***	(0.000)***	(0.01)***	(0.000)***
Growth in revenue (GR8789)	0.99	0.43	0.33	0.13
	(0.000)***	(0.000)***	(0.22)	(0.33)
Age of assets (AGE89)	-0.05	0.17	0.22	0.093
	(0.95)	(0.57)	(0.35)	(0.78)
Capital expenditure/depreciation differential (INV89)	2.65	0.98	0.69	0.93
	(0.11)	(0.11)	(2.89)	(0.13)
Import penetration (IMIO)	0.89	0.24	-0.14	-0.14
	(0.26)	(0.40)	(0.80)	(0.52)
Toxic Chemical (TRI88)	-0.03	-0.011	-0.00004	-0.00002
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Environmental Lawsuits (LAW89)	-0.00004	-0.00002	-0.00004	-0.00002
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Number of observations	233	233	233	233
Adjusted R^2	0.365	0.375	0.47	0.458

P-values are reported in (): *p < 0.10; **p < 0.05; ***p < 0.01
Industry dummy variables have been included for industries with more than seven firms in the sample (not reported here).
Models 3 and 4 also include two additional sets of interaction terms between the industry dummy variable and advertising and R&D
Source: Konar and Cohen (2001) page 288

King and Lenox (2001) also use the Tobin's Q to measure firm performance and control for similar factors as Konar and Cohen (2001) did. Furthermore, they used the TRI as environmental performance, but they went a step further and analyzed the environmental performance in and across industries. Two variables, relative emission and industry emission were designed for this purpose. King and Lenox describe these variables as follows: "Relative emission measures the firm's ability to manage and reduce its pollution by comparing the degree to which a firm's facilities are more or less polluting than other facilities in the same industry. Industry Emission measures the degree to which a firm tends to operate in the industries where production entails pollution" (King and Lenox, 2001 pp.11). As mentioned

above, they used similar control variables, but they added two further ones. Those variables were developed to take the strictness of the state law into account.

The result of their linear regression is straightforward; they find evidence for an influence of environmental performance on financial performance. By just looking at the total emission factor, it is obvious that higher emissions will lead to a decrease of firm value and the results are highly significant, as it can be observed in Table 8. Furthermore, by analyzing if firm value depends on the industry's total emission, the result shows no significance for that. But if a firm in an industry has higher emissions that its peers, they will have a significant lower firm value. Interesting is that the state law on emissions does not affect Tobin's Q.

Table 8: Estimates of future financial performance from King and Lenox (2001)

Method:	Fixed Effect 1	Fixed Effect 2	Random Effect 3	IV & Fixed Effect 4
Total Emission	-0.021**			
	(0.008)			
Relative Emission		-0.036*	-0.029 +	-0.032
		(0.028)	(0.017)	(0.021)
Industry Emission		-0.027	-0.076*	-0.083
		(0.049)	(0.037)	(0.021)
Controls				
Firm Size	-0.219***	-0.219***	-0.34*	-0.238***
	(0.030)	(0.030)	(0.014)	(0.057)
Capital Intensity	-0.420*	-0.416*	-0.147	-1.645***
	(0.198)	(0.198)	(0.187	(0.222)
Growth	0.053*	0.053*	0.068**	-0.036
	(0.022)	(0.023)	(0.022)	(0.022)
R&D Intensity	3.429***	3.377***	5.062***	1.094
	(0.535)	(0.535)	(0.408)	(0.577)
Leverage	-0.153	-0.152	-0.330***	0.149
	(0.101)	(0.101)	(0.090)	(0.110)
Regulatory Stringency	0.108	0.111	0.080*	0.035
	(0.071)	(0.071)	(0.032)	(0.107)
Permits	-0.061	-0.069	-0.060+	-0.090
	(0.045)	(0.045)	(0.032)	(0.054)
Tobin's q				-0.321***
				(0.101)
N	4483	4483	4483	3130a
Number of firms	652	652	652	544
F Stat	24.36***	22.80***		
χ^2Stat			505.30***	255.09***
Adj.R^2	0.667	0.667	0.714	0.756

The sample is slightly smaller due to the inclusion of lagged instruments.
Firm and year dummies are included but not presented in all models
Standard error are in ( )
+ p < 0.10; *p < 0.05; **p < 0.01; ***p < 0.001
Source: King and Lenox (2001), page 112

Thus both studies find significant results for TRI emissions and firm value. Furthermore, this significance is not related to a short-term market adjustment, since the performance can be observed over time as King and Lenox show in their study.

4.1.2 Quantitative Measurement of environmental factors and firm performance

In the section above, the studies were based only on one or two environmental factors, since there were no other approaches to measure environmental performance in companies. Due to an increasing interest in the environment field, a number of concepts had been developed to measure the level of environmental performance of a firm, where a various number of variables are taken into account. One of the first agencies with a quantitative environmental rating, was developed by Franklin Research and Development Corporation (FRDC). The FRDC environmental score includes the following criteria: compliance records, expenditures and other initiatives used to meet new demands, waste reduction, and support of environmental protection organizations. As one can see, the Franklin Research and Development Corporation focuses on positive components to identify environmental performance. The research agency does take certain industry specific adjustments into account and provides high and low environmental ratings for firms of all industries. Academics analyzed the correlation between the environmental criteria from FRDC and the published firm toxic releases from 1989, which were used in the above-mentioned studies. Interestingly the results show a negative correlation and they are significant (Russo and Fouts, 1997). One of the first academics that used the data from FRDC to compare environmental and stock performance, were Russo and Fouts (1997). Since FRDC only started to publish their ratings in 1991, Russo and Fouts (1997) used the data for 1991 and 1992, which is a short period. In total, they had data for 477 companies, but after controlling a number of aspects, a full dataset of 243 firms was left to analyze. They used the data from FRDC to scale the environmental score of companies. The scale went from 1 up to 5 and the highest value reflects the best environmental performance.

Russo and Fouts (1997) use a simple linear regression to observe if a firm's environmental performance had an influence on firm's return on assets (ROA). The first hypothesis from Russo and Fouts was that higher environmental performance leads to an ROA outperformance. The second hypothesis they stated was that environmental performance is more

relevant in growth industries than in slow growing ones. To test their hypotheses, they chose different control variables, which were commonly used in prior studies of performance (Russo and Fouts, 1997). The variables were firm's growth rate, advertising intensity, firm size, capital intensity, industry concentration, industry growth rate, and a dummy for pooling two years of data. The scholars generated three regressions, one to check if their control variables were correct, one to see if the environmental factor had an influence on firms return on assets, and the last one to check if the environmental factor was more relevant in growing industries.

The results are clear and as they expected. In the first model, the scholars checked their control variables. The chosen control variables were almost all significant and were in consensus with previous studies. In the second model, the focus was on the environmental performance variable, which was positive and highly significant. By comparing the value with the other control variables, one sees that the environmental rating has the highest value and therefor accounts for most of the performance. In the last model, Russo and Fouts tested the second hypothesis and besides analyzing the environmental rating, they also focus on the environmental rating in combination with growth. The results are significant and it can be said that both hypotheses were accepted, but it should be noted that all three models are supported by only a modest $R^2$of 0.30. All results can be seen in Table 9.

So over all the results show that high environmental ratings are a reason for higher stock prices, but in their robustness tests, the significance declines to a p-value of 0.11. The scholars explain that by the small sample size, which is also the basic critic on that paper, but we have to take into account that in 1997 no other data was available. In these days it's different. A number of research agencies provide data on environmental performance and some academics did a similar analysis as Russo and Fouts with longer samples, modern finance theory, and recent data. Thus the focus from research agencies shifted to different topics in ESG and that is probably the reason why the results differ so much from each other.

Table 9: Linear Regression with Environmental Rating from Russo and Fouts (1997)

Variable	Model 1	Model 2	Model 3
Constant	-2.64	-2.86	-2.99
	(2.17)	(2.16)	(2.14)
Firm growth rate	0.31**	0.30**	0.31**
	(0.03)	(0.03)	(0.03)
Advertising intensity	0.26**	0.26**	0.27**
	(0.09)	(0.08)	(0.08)
Firm size	0.49+	0.54*	0.55*
	(0.27)	(0.26)	(0.25)
Capital intensity	-0.45+	-0.50*	-0.51*
	(0.25)	(0.25)	(0.25)
Industry concentration	0.00	0.00	-0.01
	(0.02)	(0.02)	(0.02)
Industry growth rate	0.55**	0.56**	0.52**
	(0.10)	(0.10)	(0.10)
1991 dummy	2.17**	2.14**	2.27**
	(0.62)	(0.62)	(0.62)
Environmental Rating		1.49**	1.59**
		(0.52)	(0.52)
Industry growth rate*			0.51**
environmental rating			(0.18)
R^2	0.29	0.30	0.32
Difference R^2		.01	.01
F-test for difference R^2		8.26**	8.18**

N = 486. Standard error are in parentheses. Significance levels are based on two-tailed tests (F-tests are one Tailed).

+ p < 0.10; *p < 0.05; **p < 0.01

Source: Russo and Fouts (1997), page 548

Derwall, Günster, Bauer, and Koedijk (2005) for example, take the data from Innovest Strategic Value Advisors, which is more focused on risk towards different environmental facts. The research agency takes approximately 60 different key indicators into account and weights them differently according to the industry. The data is of quantitative and qualitative nature and uses over 20 information sources. The analyzed companies are compared to their industry peers and receive a rating between 1 and 10, whereby the highest score stands for high eco-efficiency. Derwall et al. (2005, p.57) summarize the sub-criteria of their research agency as follows:

- Historical liabilities: involves risk resulting from preceding actions.
- Operating risk: risk exposure from recent events.
- Sustainability and eco-efficiency risk: future risks initiated by the weakening of the company's material sources of long-term profitability and competitiveness.
- Managerial risk efficiency: the ability to handle environmental risk successfully.
- Environmentally-related strategic profit opportunities: business opportunities, such as a competitive advantage, available to the firm relative to industry peers.

They used the data over 6 years from 1997 until 2003, whereby they focused on the US market. In May 1997, their sample comprised 180 firms to form their portfolios and in May 2003 they had 450 disposable firms. They constructed two types of portfolios, one with the best eco-efficiency firms and a second with the lowest eco-efficiency scores. Derwall et al. (2005) used a linear regression to test if the environmental score generates outperformance. First they used the CAPM and found no significant results, but the high rated portfolio generated a positive alpha and the bad rated one, a negative one. To further test their approach and control for common known performance drivers they used a multifactor model. Here Derwall et al. controlled for the small-minus-big (SMB), high-minus-low (HML) and the momentum (MOM) factor. Interestingly the results show significant results for alpha as it can be seen in Table 10. Furthermore, R^2 is high and has increased compared to the tests with the CAPM. The alpha from the high rated portfolio signaled a significant average factor-adjusted return equal to 3.98% per annum and the low rated portfolio shows an underperformance of 1.08% per annum. So the study from Derwall et al. (2005) shows us, that eco efficient firms generate a significant stock outperformance. In contrast to these results, Brammer, Brooks, and Pavelin published a study one year after Derwall et al., which was based on other data and generated other results.

Table 10: Linear regression on high and low environmental portfolios from Derwall at al. (2005)

Portfolio	α_p	$(r_{m,t} - r_{f,t})$	SMB_t	HML_t	MOM_t	adj. R^2
High Ranked Companies	3.98*	0.90***	-0.22***	-0.08	-0.10***	0.87
	(1.93)	(25.02)	(-4.30)	(-1.16)	(-5.99)	
Low Ranked Companies	-1.08	0.95***	-0.15***	0.11**	-0.08***	0.88
	(-0.55)	(19.09)	(-3.70)	(2.29)	(-2.62)	
Difference	5.06*	-0.05	-0.07	-0.19**	-0.02	0.01
	(1.86)	(-0.80)	(-0.95)	(-2.20)	(-0.43)	
Industry-Adjusted Difference	6.04**	-0.20*	-0.14*	-0.30**	-0.01	0.01
	(2.38)	(-1.79)	(-1.87)	(-2.18)	(-0.18)	

T-statistics in ( ) are derived from Newey-West (1987) heteroskedasticity autocorrelation consistent standard error.
P-values are reported in (): *p < 0.10; **p < 0.05; ***p < 0.01

Source: Derwall at al. (2005), page 26

Brammer, Brooks, and Pavelin (2006) used in their work "Corporate Social Performance and Stock Returns" a similar approach to Derwall et al. As base, they also used a multifactor model and searched for a significant alpha. One of the differences between the two studies, as mentioned above, is the database. Brammer et al. (2006) used EIRIS as research agency, which defines the level of environmental performance through three factors; quality of environmental policies, environmental management system, and environmental reporting (Brammer et al., 2006 pp.101). They also give environmental specific ratings and then generate portfolios with the different scores. Since some companies do not receive ratings, they also form a portfolio with companies without any rating. The results are mixed, in fact they found a higher stock performance through a high environmental rating, but the difference between the low rated portfolio versus the high rated one is much smaller than in Derwalls work, as it can be seen in Table 11.

Table 11: Linear regression on Environmental Score by Brammer et al. (2006)

Portfolio	α_p	β_m	β_{SMB}	β_{HML}	β_{MOM}	R^2
Zero score	1.262	1.200	0.677	-0.001	-0.028	0.91
	(0.310)**	(0.075)**	0.219)**	(0.144)	(0.140)	
Tricile 1 (low score)	0.989	1.140	0.530	0.125	0.028	0.87
	(0.371)**	(0.075)**	0.282)*	(0.215)	(0.123)	
Tricile 2	1.427	1.120	0.882	0.241	0.089	0.91
	(0.327)**	(0.070)**	0.194)**	(0.138)*	(0.123)	
Tricile 3 (high score)	1.549	1.105	0.998	0.329	0.187	0.87
	(0.491)**	(0.096)**	0.231)**	(0.198)*	(0.117)	
Tricile 3 – tricile 1	0.560	-0.035	0.469	0.204	0.159	0.18
	(0.526)	(0.081)	0.239)*	(0.291)	(0.123)	
Tricile 3 – zero score	0.288	-0.095	0.320	0.329	0.214	0.19
	(0.398)	(0.109)	(0.307)	(0.189)*	(0.158)	

T-statistics in ( ) are derived from Newey-West (1987) heteroskedasticity autocorrelation consistent standard error.
P-values are reported in (): *p < 0.10; **p < 0.05; ***p < 0.01
R^2 is for the 4-factor regression
Source: Brammer et al. (2006), page 111

Overall, we can summarize that environmental performance seems to be relevant for firms financial performance, but it has to be mentioned that the style of how environmental performance is measured, has to be taken into account because it's crucial.

4.2 Social component in ESG Investments

The social aspect these days is clearly related to the employee and other stakeholders of the firm. Interesting is that by reading through the traditional theory of business management (e.g. Taylor, 1911; Friedman, 1970), the employee had no status in the company and could be seen equally important as raw materials. The goal of the management was to be as efficient as possible by keeping the costs on the lowest level as possible (Edmans, 2009). Especially from the 19th to mid-20th century, companies used simple assembly lines for their mass production, so that workers were able to do their specific tasks one after another. Frederick Taylor (1911) went even one step further in his work "The Principles of Scientific Management". He stated that the management should exactly define the work tasks and the time horizon for each workman in advance (Frederick Taylor, 1911). To capture his concept in other words, it can be said that workers were paid to tackle physical tasks and not for thinking.

In opposite to the traditional theory, recent theories argue that the role of human capital has dramatically changed in the last 50 years. In many studies the employee has become the heart of the company, since the modern company driver is quality and innovation (Frederick Herzberg, 1959; Douglas McGregor, 1960), rather than simple raw material, which can be bought and sold easily. For example, a study from McKinsey & Company, where they constructed a survey with over 6000 managers concluded that talent would be crucial in the next 20 years for the evolution of companies (Fishman, 1998). The idea is that employees can invent new products and strengthen client relationship, from which a company can profit in the future, but where does the motivation come from? Akerlof and Yellen (1986) stated in their work about the efficient wage theory that satisfaction can increase effort because workmen do not want to get fired from a satisfying job, so they will work harder (Shapiro and Stiglitz, 1984) and if workers receive an additional compensation, they will see it as a gift and increase their productivity to thank the corporation (Akerlof, 1982). Furthermore, some studies state that a good social environment reduces costs of lawsuits (Brammer et al., 2006; Ullman, 1985).

Above we mentioned motivation and job satisfaction, which can be achieved by the firm through a high level of compensation, but this doesn't necessarily mean that a high compensation is achieved with higher paychecks. In the past, cash was seen as the most effective

motivator for work, since workmen needed cash to satisfy their physical needs (Edmans, 2009). Recent studies yielded that a high paycheck only motivates up to a certain point (Herzberg, 1959) because as soon as the physical needs are met the workers satisfaction increases through recognition and self-esteem. In consensus with the newer theory, the satisfaction of employees seems to play a bigger role for the workmen willingness to perform well, but the question occurs, if higher employee satisfaction and motivation leads to a higher output and this further implies a higher company performance. To receive an answer to this question, the first step is to identify the compensation and benefits a workman receives from his employer.

It is very difficult to measure satisfaction of an employee, since so many factors can be relevant and each worker is satisfied by different components. This includes work-life balance, compensation, stock plans, alternation, recognition, self-esteem, occupation, and other benefits, which are designed for the employee. With all the single elements, the question arises how the social aspect can be evaluated over a large quantity of firms.

Past studies mainly use quantitative databases from e.g. CEP and KLD, which gather their data through public information and engagement with companies, but they do not have direct contact with employees of the companies to identify the satisfaction. So, simple measurements like minority representation have a strong impact on the social rating, which can manipulate the results and could be the reason for the insignificant findings (Edmans, 2009). This is why more recent studies tried to find a better measurement of how to capture employees' satisfaction. To understand the different approaches, we looked at a number of studies to receive a better understanding. We split this part of the study into a quantitative and qualitative section.

4.2.1 Quantitative approach to capture Employees Job Satisfaction

Brammer, Brooks, and Pavelin (2006) are a group of academics, which try to measure the impact of corporate social performance on stock returns through a quantitative approach in their work "Corporate Social Performance and Stock Returns". In their paper, they analyze community, environmental, and employee performance and link these factors to stock return. Since this section of the paper focuses on the social (employees) performance, there

will be a focus on the social section of their paper. A part of this paper was already addressed in the environmental section.

Brammer et al. (2006) use the Ethical Investment Research Service (EIRIS) as database, which evaluates firms' corporate social performance by using an "objective" set of criteria (e.g. amount of environmental fines, women working in the board, etc.) The provided data is mixed, some components are quantitative and others are in text. To provide useful information, EIRIS updates the company data on a continuous basis. The data for each corporation is updated at least twice a year and if any extraordinary event occurs the research agent implements these events whenever they happen. This approach makes the distribution constantly comparable and fair (Brammer et al., 2006). Brammer et al. used the data provided in February 2003, which received their last update 5 months earlier. The indicator employee, in which we are interested in, consists of six different components relating to health and safety, employee relations, employee training and development, equal opportunities system and policies, and systems for job creation and security (Brammer et al., 2006). These components were then transformed into quantitative measurements. Each element receives a score between 0 and 3, so that in total the employee responsibility score can reach up to 18 points. 44% of all companies had no score, which does not imply bad environmental performance. The explanation of Brammer et al. is the so called small firm effect. Small companies do not have the financial flexibility to implement corporate social responsibility (CSR) reports to receive an evaluation form EIRIS. During their research, the academics found out: firms in an industry, where a high exposure to a certain factor can be identified, do engage and report more about it, to lower their reputation risk. The thereby generated result is that these firms receive a higher score. So it seems that the collected data is a subjective decision of the analyst/reader of the companies CSR report.

As first step they produce various portfolio returns, formed out of the different CSR scores, over a 1-, 2-, and 3-year period. The result shows a negative performance for all components in the first year, which is linked to a globally bad stock performance, but the companies with high employment/social score performed better than companies with low scores, as shown in Table 12. In the following two years, market performance was positive and the results are mixed. In the second year, bad and good environmental firms perform similar and in the third year, the portfolio with low scores outperformed the other ones.

Table 12: Yearly Social Portfolio Performance from Brammer et al. (2006)

Portfolio	1-Year Return	2-Year Return	3-Year Return	n
Zero score	-1.42	-0.26	0.14	28
	(2.61)	(1.59)	(1.27)	
Tricile 1 (low score)	-0.95	0.52	0.72	90
	(3.18)	(1.55)	(1.19)	
Tricile 2	-1.04	0.12	0.33	89
	(2.83)	(1.97)	(1.47)	
Tricile 3 (high score)	-0.46	0.52	0.55	89
	(2.41)	(2.02)	(1.56)	

Standard deviations of returns over time in ( )
Results are based on equally weighted portfolios
Source: Brammer et al. (2006), page 107

Furthermore, Brammer et al. used the Carhart Model to find out where the performance could come from. The Carhart Model is based on the three-factor model and adds a fourth momentum factor to the model. The formula is represented in Equation 8. By regressing the different employment portfolios on the Carhart Model a significant alpha had been observed. It is interesting that all portfolios, from the lowest employment score up to the highest, generate a positive alpha. However, the best-scored portfolio also has the highest alpha, as it can be seen in Table 13.

$$r_{p,t} = \alpha_p + \beta_m\left(r_{m,t} - r_{f,t}\right) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) + \beta_{MOM}(MOM_t) + \varepsilon_t \qquad [9]$$

Table 13: Carhart Model on portfolio based on social performance from Brammer et al. (2006)

Portfolio	α_p	β_m	β_{SMB}	β_{HML}	β_{MOM}	R^2
Zero score	1.167	1.231	1.069	0.019	-0.158	0.80
	(0.550)**	(0.106)**	(0.406)**	(0.332)	(0.242)	
Tricile 1 (low score)	1.385	1.225	0.591	0.086	0.103	0.90
	(0.342)**	(0.065)**	(0.243)**	(0.143)	(0.119)	
Tricile 2	0.938	0.996	0.612	0.142	-0.031	0.94
	(0.229)**	(0.047)**	(0.142)**	(0.103)	(0.081)	
Tricile 3 (high score)	1.606	1.160	1.010	0.356	0.193	0.88
	(0.445)**	(0.093)**	(0.241)**	(0.195)*	(0.122)	
Tricile 3 – tricile 1	0.221	-0.064	0.419	0.270	0.090	0.30
	(0.394)	(0.069)	(0.187)	(0.199)	(0.103)	
Tricile 3 – zero score	0.439	-0.071	-0.059	0.338	0.351	0.12
	(0.553)	(0.115)	(0.521)	(0.327)	(0.266)	

T-statistics in ( ) are derived from Newey-West (1987) heteroskedasticity autocorrelation consistent standard error.
P-values are reported in (): *p < 0.10; **p < 0.05; ***p < 0.01
R^2 is for the 4-factor regression

Source: Brammer et al. (2006), page 111

4.2.2 Qualitative approach to capture Employees Job Satisfaction

Edmans (2009) chose a different approach to measure the employees' satisfaction, since he believes that the quantitative approximation is too subjective. That's why he used the list "100 Best Companies to Work for in America" for a qualitative approach, because he sees a closer linkage between the results and employees. A number of studies in popular press (e.g. Branch, 1999; Grant, 1998; Lau & May, 1998) used the same approach and the same database, but they only focused on simple benchmarks and did not have such a long data series as Edmans, but almost all of them found positive results. The "Top 100" list was first published in 1984 by Katz, Levering, and Moskowitz and was updated the first time nine years later by two of the inventors. 1998 the *Fortune Magazine* decided to publish, in cooperation with Levering and Moskowitz, the List each year in January.

Companies have to apply to be considered for the list. This self-selection is reasonable since firms with internal social difficulties would not apply and the list is internationally well known so the author expects that all interested firms will send their application to the "Great Place to Work, Institute of San Francisco". The application has to take place one year in advance. The evaluation is generated through two valuation types: one third of the score comes from observable company factors (e.g. demographic makeup, culture, paychecks, and other benefit programs). The other two thirds are allocated to an employee survey, which includes approximately 60 questions concerning management, camaraderie, job satisfaction, and fairness. This questionnaire is send to 250 randomly selected employees in each firm. After receiving all the relevant information, companies receive scores in four different sections: employee communication, opportunities/benefits, fairness/compensation, and pride/camaraderie.

The research from Edmans concentrates on the above-mentioned list. For each year the data was published (1984, 1993, and 1998-2005), he created a portfolio with all publicly traded companies from the 100 Best Companies List. So the author only focuses on investing in "good" firms and doesn't consider generating a short portfolio for the "bad" firms. He compiles both, equal- and value-weighted portfolio because Fama and French (2008) found out that a number of anomalies are not robust to the weighting methodology (Edmans, 2009). After creating the portfolios, Edmans makes sure that the generated performance from companies of the Best Companies List does not simply come from the common known risk factors like small cap, growth firms or from the momentum effect. To adjust his research for

the factors, Edmans used, similar to Brammer et al., the Carhart model (1997) as shown in Equation 8. Edmans just changed the order of the variables.

Contrary to Brammer et al. (2006), Edmans uses three different benchmarks to have an optimal benchmark for each request. The first benchmark is simply the risk-free rate, which will not provide us with redounding outputs. The second benchmark is an industry-matched portfolio designed by Fama and French (1997), to ensure that the performance doesn't come from outperforming industries. The last benchmark used is a characteristic-adjusted benchmark build by Wermers (2004)[7]. The third benchmark seems to be the most effective one, since it focuses on the risk factors from the Carhart model.

For us, alpha is again the most interesting component in the equations because its represents performance through stock selection and the stock selection is allocated to the performance generated through selecting "Top 100 List" as assets. The other factors are identical to the factors described in our methodology and the before mentioned studies.

Edmans wanted to find out if employee satisfaction can be linked to stock performance and if the market is able to fully incorporate this association. The result is straightforward and consistent over time: over the 22 years of data, the portfolio outperformed the market in 18 years (Edmans, 2009) and this is consistent over all three benchmarks and portfolios. Clearly, the highest outperformance is generated over the risk-free rate, but also by analyzing the results over the other benchmarks, the social portfolio generates an outperformance of 0.19% - 0.25% per month as it can be seen in Table 14. This is an outperformance of 2.28% - 3% per annum. Even during the years between 2000 and 2002, when the market saw a strong decline, the portfolio outperformed the market (Edmans, 2009). All alpha values are significant and show that investing into "The Best Companies to Work for" will generate outperformance. In Table 14 the results from Edmans (2009) can be found in more detail. Panel A represents the results for the equal-weighted portfolio and Panel B shows Edmans results for the value weighted portfolio.

[7] Benchmark under : http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm

Table 14: Carhart Model on the Top 100 List from Edmans (2009)

	Excess return over				Excess return over		
	Risk-free	Industry	Characteristics		Risk-free	Industry	Characteristics
Panel A (equal-weighted)				Panel B (value-weighted)			
α_p	0.34	0.22	0.25	α_p	0.34	0.20	0.19
	(3.49***)	(2.97***)	(2.97***)		(3.03***)	(2.70***)	(2.63***)
β_m	1,11	0.07	0.12	β_m	0.95	-0.03	-0.01
	(38.08***)	(3.41***)	(4.74***)		(30.29***)	(1.30)	(0.27)
β_{HML}	0.03	0.04	0.02	β_{HML}	-0.46	-0.09	-0.15
	(0.64)	(1.23)	(0.61)		(8.13***)	(2.39***)	(3.77***)
β_{SMB}	0.15	0.14	0.03	β_{SMB}	-0.24	-0.25	-0.04
	(3.08***)	(4.45***)	(0.80)		(4.77***)	(7.23***)	(1.27)
β_{MOM}	-0.13	-0.04	-0.08	β_{MOM}	-0.04	-0.00	-0.02
	(4.76***)	(2.20**)	(3.97***)		(0.95)	(0.06)	(1.03)
# obs	262	262	262	# obs	262	262	262

t-statistics are in ( ) and are derived from Newey-West (1987) heteroskedasticity autocorrelation consistent standard error
P-values are reported in: *p < 0.10; **p < 0.05; ***p < 0.01
Source: Edmans, 2009, pp. 26

Edmans recognizes that the above mentioned outperformance could be a result of missing public information, since until 1998 the companies were only available in books not open for public. That's why he repeated the study for a sub-period from 1998 to 2005. During this period, the Fortune Magazine published the list on an annual basis. The outperformance is even higher and still significant. During the sub-period, Edmans found a monthly outperformance of 0.32% - 0.64% as shown in Table 15. Here again, Panel A shows the results for the equal-weighted portfolio and Panel B for the value-weighted portfolio.

Table 15: Carhart Model (98-05) on the Top 100 List from Edmans (2009)

	Excess return over		
	Risk-free	Industry	Characteristics
Panel A (equal-weighted)			
α_p	0.64	0.46	0.57
	(3.70***)	(3.28***)	(4.08***)
Panel B (value-weighted)			
α_p	0.47	0.30	0.32
	(2.06**)	(2.05**)	(2.11**)
# obs	96	96	96

T-statistics in ( ) are derived from Newey-West (1987) heteroskedasticity autocorrelation consistent standard error.
P-values are reported in (): *p < 0.10; **p < 0.05; ***p < 0.01

Source: Edmans, 2009, pp. 27

As most of the other studies, Edmans only focused on the performance and did not analyze any other factors, which could provide further relevant information. In contrast to Edmans, an earlier study from Fulmer, Gerhard, and Scott (2003) examined the data used in the "Top 100 List" survey and also looked at the return on assets (ROA) and the market-to-book (MTB) value of firms. Fulmer et al. (2003) state the hypothesis, that both values should be on average higher for the Top 100 firms because positive employment relations lead to higher customer satisfaction (e.g. Koys, 2001) and this leads to a constant or increasing revenue stream (e.g. it is more expensive to acquire new costumers then to keep old ones). Furthermore, Fulmar et al. (2003) mention that higher employee satisfaction would attract a larger number of applicants for jobs, so that companies are able to select better-qualified and motivated employees. They expect a higher market-to-book because the market allocates a higher value to the intangible assets, which would include satisfied and motivated workers.

Table 16: KPI Comparison with a social aspect from Fulmer, Gerhard, and Scott (2003)

	Portfolio median -> return on assets			Portfolio median -> market-to-book ratio		
Year	100 Best	Matched Firms	Z statistic	100 Best	Matched Firms	Z statistic
1995	9.86%	7.11%	0.41	3.69	2.64	0.92
1996	10.09%	8.17%	0.96	4.20	3.05	1.57
1997	10.78%	7.76%	3.23**	4.91	3.18	2.49*
1998	9.03%	6.65%	2.66**	4.34	2.32	2.62**
1999	9.15%	7.64%	1.76+	4.43	2.77	2.23*
2000	7.93%	6.96%	1.68+	5.40	2.17	2.34*

Sample size range from 43 to 50 for ROA and
39 to 48 for market to book
+ p < 0.10; *p < 0.05; **p < 0.01
Source: Fulmer, Gerhard, and Scott (2003), page 983

In Table 16, the results from Fulmer et al. (2003) are apparent and it seems that their hypothesis is true. In the period from 1995 to 2000, higher ROAs and higher MTBs occur. As mentioned before 1998 was the first year where the list was published in the Fortune magazine, and from there, the results were always significant. They believe that the publication enabled investors to detect intangible values (satisfied employees). In 1997, Fulmer et al. (2003) also found a highly significant result for the difference between the Top 100 and matched firms, but they did not find any explanation for that. The largest difference between the List and the benchmark appears in 1997 with 10.78% to 7.76% and the value from the z-statistic is with 3.23 highly significant. In the years 1995 and 1996, there is a difference between the ROA, but the z-statistics are only 0.41 and 0.96, which indicate no significance. Fulmer et al. (2003) explained it with missing public availability of the list.

On the right side of the table we see the results of the market-to-book-ratios and here, the values are also significant from 1997 on. The highest significance can be seen in 1998, when the List was published for the first time in the Fortune magazine. The difference between the Top 100 List and the benchmark is 2.02. Furthermore, they also analyzed the performance between firms from the “Top 100 List” and a benchmark. As well as Edmans, they found significant outperformance for the Top 100 Companies, but their sample period is quite short.

Overall, the qualitative approach from Edmans (2009) found a significant outperformance over all sub-periods and benchmarks. Furthermore, an earlier study analyzed the ROA and MTB, which could indicate that the list is used from investors to identify the intangible asset “employees’ satisfaction” in firms. On the quantitative side, there was some head-to-head outperformance but the analyzed study was not able to find real and consistence significance. This also supports the hypothesis from Edmans, that a quantitative approach is not representative for the social factor, since the data comes from the corporation and not directly from the employee.

4.3 Corporate Governance in ESG Investments

Corporate governance can be compared to a republic. In the story of Robin Hood a good king, Richard Lionheart, who acts in the interests of his subjects, hands over his kingdom to his evil brother, John Lackland, who steals and plunders as much as he can, much to the disadvantage of his subjects. At the end this provokes a rebellion led by Robin Hood to restore peace and order and Richard’s rule. What sounds like a children’s story is nothing less than an archetypical description of one of economics’ most basic dilemmas: the principal-agent problem.

Such a problem arises whenever “two parties have different interests and asymmetric information (the agent having more information), such that the principal cannot directly ensure that the agent is always acting in his (the principal's) best interests” (Hanf, Iselborn and Jungbluth, 1999), particularly when activities that are useful to the principal are costly to the agent, and where elements of what the agent does are costly for the principal to observe.

One of the most important examples of this dilemma manifests itself in the management of the modern joint stock company where ownership and management are separated. Share-

holders (principals) hand over control to corporate management and hope that it will act in their best interest. Nowadays, the role Robin Hood used to play is played by Corporate Governance, which includes the board of the company. Checks and balances, rules and regulations that make sure that corporate management does not steal and plunder, but acts in the interests of their principals.

There are many different definitions of corporate governance, but they all fall within close proximity of each other. We prefer the definition established by Kaplan and Schweser from 2012.

“Corporate governance is the set of internal controls, processes, and procedures by which firms are managed. It defines the appropriate rights, roles, and responsibilities of management, the board of directors, and shareholders within an organization. It is the firm's checks and balances. So we can say that a company is well governed if the board of directors protects shareholder interests.”

Before the 1980s, large companies had almost no reason to restrict shareholder or management rights. Investors’ interest in taking on an active role in the company was almost non-existent until then, and hostile takeovers and proxy fights were rare events. The appearance of the high-yield bond market in the 1980s facilitated hostile takeover offers for even the largest public companies (Gompers, Ishii and Metrick, 2003). During this period the importance of instruments under the umbrella of, and available to, the corporate governance term experienced an enormous hype. As soon as the market, including the companies themselves, realized the power of corporate governance tools, many academics, regulators, and even investors concluded that something was "wrong" with the way the companies were governed. They all realized that the scope of actions available to company management was excessively large. The emergence of this issue triggered discussions about possible solutions. In addition to the corporate governance theme, some critics went even further and suggested to make companies more socially responsible (Fischel, 1982). After a short period of time the US government reacted. Bills with an enormous number of reform proposals were being presented to the US Congress (Corporate Democracy Act of 1980). In 1980 the Securities and Exchange Commission published an exhaustive document entitled "Staff Report on Corporate Accountability” - discussing the perceived problem and many of the proposed solutions. But all of this was just the beginning of corporate governance. Over

time, more and more investors got interested in the topic and realized that they could use their shares to influence the companies and increase its value (Fischel, 1982).

Corporate Governance is the oldest topic in the ESG field, but still young. It only came up in the 1980th, and had similar difficulties in measuring and defining the issues in the field. So we will follow a time line in this dimension to show how corporate governance has evolved. We will focus on papers that analyse how and why weak corporate governance affects companies, their share prices, and other corporate characteristics. Whereby we will focus on a crucial paper, which was written by Gompers, Ishii, and Metrick (2003). They defined a broadly accepted concept of how to measure corporate governance, which is still mainly used.

4.3.1 Early Measurement of Corporate Governance

The field of corporate governance offers many topics for investigation, but one of the crucial questions is how to measure corporate governance. Before 2003, there were many different approaches. For example in 1998 Core, Holthausen, and Larcker analyzed if corporate governance influenced company performance. They stated the hypothesis that weak corporate governance led to higher agency costs for shareholders and that this in turn led to bad share price performance. But how did they determine the level of corporate governance in a company? In their work Core et al. (1998) only focused on one factor: the compensation of the CEO to determine corporate governance. But this factor was influenced by many different forces. In particular, they concentrated on board and ownership structures and found that CEOs earned higher compensation when governance structures were less strict. With respect to ownership structure they showed that an increase in CEO equity ownership resulted in a reduction of CEO compensation. Also, excess CEO compensation was generated by board size and the percentage of the board made up by external directors appointed by the CEO, the number of board members older than 69, external directors who served on several boards, and whether the board chair was also the CEO (Core et al., 1998).

The question that is actually interesting for us is whether weak corporate governance affects stock performance. After Core et al. had cleared the data and defined excessive CEO compensation, they designed a linear regression to determine the impact of excessive compensation on share price return.

The results were clear; the linear regression indicated that excessive CEO compensation had a significantly negative effect on the company's share price return (another linear regression in the paper also showed a negative correlation with the company's performance). By looking at the results we can establish, that an increase of 100% of CEO compensation will approximately lead to a 12.432% decrease in the share price in the first year. The influence of higher compensation decreases over time, but remains negative.

Table 17: Regressions of subsequent stock return on predicted excess CEO compensation from Core et al. (1999)

	Average return for period		
	One-year	Three-year	Five-year
Predicted excess compensation	-12.432	-7.059	-4.438
	(-3.73)	(-3.47)	(-2.54)
σ_p	-0.205	-0.226	-0.194
	(-2.70)	(-4.98)	(-5.00)
ln(Market value of equity)	-3.666	-0.656	-0.464
	(-2.95)	(-0.87)	(-0.72)
Market-to-book ratio	-9.446	-4.489	-3.037
	(-4.55)	(-3.59)	(-2.79)
Adj-R^2	30.6%	28.2%	31.9%
F	12.39	10.5	11.29
Number of regressors	20	20	20
n	491	460	418

Standard deviation of return is based in the annual percentage of stock market return
T-statistic are based on OLS standard errors and are shown in ( )
Source: Core et al., 1999, p.395

4.3.2 New Measurement of Corporate Governance ("Governance Index")

In 2003 Gompers, Ishii, and Metrick (2003) defined a broadly accepted concept of how to measure corporate governance of a company by introducing the so-called "Governance Index". They established the index in order to find an appropriate measure for the level of corporate governance. And lastly they examined whether a certain level of corporate governance had an impact on key figures of a company, such as share price return, company value, profits, or sales growth.

Each company consists of three different parties: shareholders, directors, and managers. Shareholders are allowed to vote and consequently elect directors. These directors most commonly delegate their decision-making and responsibilities to managers. Within such a system, rules and regulations allocate the level of power to each of these parties. Considering two extreme outcomes one could be labelled "democracy", in which shareholders can enjoy a high level of rights, especially the right to vote and consequently; here, directors can be replaced quickly and easily. At the other end, "dictatorship", where the main power is in the hand of managers and shareholders' rights are strongly restricted. Gompers, Ishii, and Metrick used this kind of extremes in order to find whether there is a relationship between shareholder rights, resulting in a certain level of corporate governance and corporate performance (Gompers et al., 2003).

The index consists of 24 corporate governance provisions, including company-level provisions and state laws. The index was built by analyzing 1,500 firms from 1990 to 1999. These 24 distinct provisions are grouped into five thematic groups:

Delay: including all tactics for delaying hostile bidders *(included components: Blank Check, Classified Board, Special Meeting, Written Consent)*

Protection: implying different arrangements in order to protect officers and directors *(included components: Compensation Plans, Contracts, Golden Parachutes, Indemnification, Liability, Severance)*

Voting: provisions focusing on shareholders' rights in elections or charter/bylaw amendments *(included components: Bylaws, Charter, Cumulative Voting, Secret Ballot, Supermajority, Unequal Voting)*

State: since the company-level data based on the data source IRRC do not automatically entail all these provisions to fall under state law, this last group is added *(included components: anti-greenmail Law, Business Combination Law, Cash-Out Law, Directors' Duties Law, Fair Price Law, Control Share Acquisition Law)*

Other: additional company-level provisions *(included components: Anti-greenmail, Directors' Duties, Fair Price, Pension Parachutes, Poison Pill, Silver Parachutes)*

The index construction is very easy and straightforward; one point is added for every provision that reduces shareholder rights. A higher score implies more restrictions on shareholder rights. Going back to the concept of comparing companies with political systems, this system leads to so-called "Democratic Portfolios" – firms in the lowest decile of the index, representing lowest management power and strongest shareholder rights – and "Dictatorship Portfolios" – firms in the highest decile with exactly the opposite power-sharing relationship (Gompers et al., 2003).

A closer look at the five groups reveals that coverage under state laws is not highly correlated with the adoption of company-level provisions, whereas the company-level sub-indices show a significant positive correlation among themselves. Furthermore, companies with a high level of corporate governance have certain characteristics. Strongly governed companies are usually small companies, which are not included in the S&P index. Another important insight is that well-governed companies do not tend to decrease their level of governance over time. Of the 10 best firms in the Democracy Portfolio in 1990, six kept the lowest G score, and the rest only dropped slightly in their level of corporate governance (Gompers et al., 2003).

As the next step, Gompers et al. had a closer look at the relationship between the G index and subsequent returns. The results, represented in Table 18, show that the Democratic Portfolio outperformed the Dictatorship Portfolio by more than 8.5% per year over the entire period. In closer detail, an investment of USD 1 in the Dictatorship PF on the 1st of September 1990 would be worth USD 3.39 nine years later, whereas the same investment in the Democratic PF would have yielded a gain of USD 7.07 over the same period (Gompers et al. 2003, p. 121). To ensure that parts of the outperformance were generated by good corporate governance, the scholars tried to find out the reasons behind this disparity with the help of the four-factor model of Carhart, (1997), which is shown in Equation 9:

$$r_{p,t} = \alpha_p + \beta_m\left(r_{m,t} - r_{f,t}\right) + \beta_{SMB}(SMB_t) + \beta_{HML}(HML_t) + \beta_{MOM}(MOM_t) + \varepsilon_t \qquad [10]$$

Variables are similar to the ones describes in the earlier parts of our work. $r_{p,t}$ is defined as the excess portfolio return in month *t*, $r_{m,t}$ is the value-weighted market return, $r_{f,t}$ is the risk-free rate, the terms *SMB* (small – big), *HML* (high – low) and MOM stand for the month *t* returns in order to capture size, book-to-market, and momentum effects, respectively.

Again, the most relevant intercept for us is α_p since it takes the performance through selection into account. As seen in Table 18, this method results in an alpha of 71bps per month, which equals an annual 8.52% (buying strong and selling weak CG firms). This means that the usual outperformance factors are not the only reason for the outperformance. Upon closer scrutiny we find that the Democracy PF outperforms the Dictatorship Portfolio significantly, earning the highest alpha of all the deciles (G5-G14) (Table 18).

Table 18: Governance and Returns: Performance-Attribution Regressions for Decile Portfolio from Gomper et al. (2006)

	α_p	$(r_{m,t} - r_{f,t})$	β_{SMB}	β_{HML}	β_{MOM}
Democracy-Dictatorship	**0.71****	**-0.04**	**-0.22***	**-0.55***	**-0.01**
	(0.26)	**(0,07)**	**(0.09)**	**(0.10)**	**(0.07)**
G≤5 (Democracy)	0.29*	0.98**	-0.24**	-0.21**	-0.05
	(0.13)	(0.04)	(0.05)	(0.05)	(0.05)
G=6	0.22	0.99**	-0.18**	0.05	-0.08
	(0.18)	(0.05)	(0.06)	(0.07)	(0.04)
G=7	0.24	1.05**	-0.1	-0.14	0.15**
	(0.19)	(0.05)	(0.07)	(0.08)	(0.05)
G=8	0.08	1.02**	-0.04	-0.08	0.01
	(0.14)	(0.04)	(0.05)	(0.06)	(0.04)
G=9	-0.02	0.97**	-0.20**	0.14**	-0.01
	(0.12)	(0.03)	(0.04)	(0.05)	(0.03)
G=10	0.03	0.95**	-0.17**	0	-0.08**
	(0.11)	(0.03)	(0.04)	(0.04)	(0.03)
G=11	0.18	0.99**	-0.14*	-0.06	-0.01
	(0.16)	(0.05)	(0.05)	(0.06)	(0.04)
G=12	-0.25	1.00**	-0.11*	0.16**	0.02
	(0.14)	(0.04)	(0.05)	(0.06)	(0.04)
G=13	-0.01	1.03**	-0.21**	0.14*	-0.08*
	(0.14)	(0.04)	(0.05)	(0.06)	(0.04)
G≥14	-0.42*	1.03**	-0.02	0.34**	-0.05
	(0.19)	(0.05)	(0.06)	(0.07)	(0.05)

Standard errors are represented in ( )
*p < 0.05; **p < 0.01
Source: Gompers et al. 2003, page 123

So it can be said that higher corporate governance indicates higher share prices, but due to the short sample period one could potentially argue that there was a period of mispriced company values. This argument can be eliminated by analysing the company value for different corporate governance levels. This method is based on an regression model by Kaplan and Zingales (1997), using the Tobin's Q as valuation measure. The independent variables are a vector of governance variables (G or modifications of G) and a vector of company characteristics. The results are straightforward; a higher level of G (lower shareholder rights) implies a reduction of Q (company value). The largest absolute value occurs in 1999 with a significant coefficient of -0.114, which means a one-point increase in G, equivalent to adding a single governance provision, is associated with decrease of 11.4 percentage points in Q" (Gompers et al. 2003, p.18). The results illustrate that when the sample is restricted to companies in Democracy and Dictatorship PF the results are the same. Companies in the Democracy PF show greater value. Looking at the estimated point in 1999 again, we find that the implication is even stronger, since the coefficient is 0.563, which translates into a 56% higher Q value for companies in a Democracy PF (Gompers et al., 2003).

Furthermore, the study of Gompers et al. includes results of regressions for net profit margin, return on equity, and sales growth with G as the key regressor. All average coefficients on governance are negative but the regression is significant only for the net margin and sales growth. Overall it can be concluded that a higher G, which implies lower shareholder rights, has a negative impact on operating performance on a firm (Gompers et al., 2003).

4.3.3 Modification of the G-index: The E-index

In 2010, Cremers and Ferrell extended the research in the field of corporate governance. They based their study on the already explained and widely used G index, but expanded the time period from 1978 to 2006 for 1,000 companies. They also included in their study the E index of Bebchuk, Cohen, and Ferrel (2009), which is based on six provisions of the G index (supermajority voting requirements for charters, bylaws and merger, classified boards, poison pills, and golden parachutes).

Some of their results are in line with the earlier stated outcomes by Gompers et al. Again, Cremers and Ferrell examined if there was a relationship between the G index (and E index)

and Tobin's Q. Their results confirmed that a higher G index implied a lower Tobin's Q associated with a lower company value. This finding was also valid for the E index (Cremers and Ferrel, 2010).

Furthermore, they tested whether any evidence could be found in support of the "reverse causation" explanation, e.g. firms with higher value tended to adopt more shareholder rights or via versa (Cremers and Ferrel, 2010). But they found only limited support for this rationale. As long as they did not include any firm and effects fixed to a particular year, they would find a positive relation, meaning that companies with higher values tended to adopt more G index provisions. But this significantly positive association collapsed when fixed effects were added (Cremers and Ferrel, 2010).

4.3.4 The Corporate Governance Outperformance Drivers

After introducing a diversified empirical research base we can conclude that well-governed firms perform better than companies with weak corporate governance. This prompts the question: where does the performance come from? Many researchers, including Gompers et al. and Core, Guay, and Rusticus, tried to find an explanation for the correlation of the G index with returns, company value, and operating performance. First of all, Gompers et al. (2003) formulated three different hypotheses that might explain the results discussed earlier. They are stated as followed:

Hypothesis I: Governance provisions cause higher agency costs. These higher costs were underestimated by investors in 1990.

Hypothesis II: Governance provisions do not cause higher agency costs, but rather were put in place by 1980s managers who forecast performance for their companies in the 1990s.

Hypothesis III: Governance provisions do not cause higher agency costs, but their presence is correlated with other characteristics that earned abnormal returns in the 1990s.

According to hypothesis I, researchers state that a reduction in shareholder rights have an impact on rising agency costs which indirectly linked to some findings of Jensen (1968). Inefficient investment decisions or the extraction of private benefits by managers, aided by the fact that it was more difficult or costly for shareholders to replace managers, might be the reason for increasing agency costs. Examining this hypothesis they looked at capital expendi-

ture on the one hand and acquisition activity on the other hand. For capital expenditure they argued that managers might undertake inefficient projects to generate private benefits, and even more so in cases where they were trying to thwart hostile takeovers. Consequently an increase in corporate CAPEX would be in line with the adoption of new takeover defences and would ultimately cause a decrease in company value. The result showed that the average coefficient on governance was significantly positive, which led to the conclusion that firms with lower shareholder rights (high governance) had higher CAPEX than low governed companies. Looking at the relation between acquisition activity (number of acquisitions) and governance, the average coefficients showed positive and significant results, meaning that high governed companies might have engaged in high volumes of inefficient investments during the 1990s (Gompers et al., 2003).

Another study, which can be linked to hypothesis I is the work of Core, Holthausen, and Larcker (1998) where they found that bad share price performance came with higher CEO compensation, as manifesting itself in agency costs and bad corporate governance. Dittmar and Mahrt-Smith (2007) went further and looked for the actual reason. One part of their paper concentrated on the allocation of cash holdings and the level of corporate governance in companies. They argued that the usage of cash was the simplest way to identify an agency problem in companies. In an effort to prove this assumption, Dittmar and Mahrt-Smith used a study of Faulkender and Wang (2005), which focused on the potential benefits of holding cash, showing that the value of one dollar of cash was often less than one. In their study they found that USD 0.94 translated into one dollar held in cash (Faulkender and Wang, 2005).

Dittmar and Mahrt-Smith (2007) went further by examining the cost of holding cash and relating the low value of cash to the company's corporate governance. To do so, they used two different approaches to identify the level of governance. One approach was the G-Index that we have mentioned earlier in the paper; the other one was to investigate how influential the biggest shareholders were in the company. They then proceeded to examine how excess cash of a firm was valued according to different levels of corporate governance. In addition they analyzed how much value is attributed to each additional dollar in cash. Both studies showed that, if a firm was well-governed, the value of cash was higher and would

increase the firm value. In Table 19 it can be seen that the marginal value of one dollar is USD 1.62 in a well-governed company and is only USD 0.42 in a badly governed one.

By looking at all the different studies that focused on hypothesis I or could be linked to it, we find that the outperformance can be attributed to lower agency costs in a strongly governed company, since the wealth of shareholders will be maximized.

Table 19: The marginal value of cash for the average company from Dittmar and Mahrt-Smith (2007)

Sample means for cash value	Lagged Cash	0.12	0.12
	Leverage	0.22	0.23
	Gompers; Ishii and Merick Index Dummy	0.56	
	Blockholdings Dummy		0.51
Marginal Value of USD 1 (Good Governance)		**1.62**	**1.27**
Marginal Value of USD 1 (Poor Governance)		**0.42**	**0.88**
Marginal Value of USD 1 (Average Governance Firm)		**1.09**	**1.07**

Source: Dittmar and Mahrt-Smith, 2007,page 611

For the second hypothesis, which referred to insider information, Gompers et al. were not able to find statistical evidence. They were basing this hypothesis on the idea that managers in the 1980s predicted poor performance, but investors did not have this information. Consequently managers increased governance provisions in order to secure their jobs. Although the provisions might have real protective power, they would not have been the reason for the poor performance. Due to the lack of evidence they had to reject this explanation (Core et al., 2006).

The third explanation, Hypothesis III, was associated with risk or other firm characteristics, like size, price, membership in the S&P 500, the dividend ratio, and some additional factors that caused the abnormal returns in the 1990s. Only a small portion of these different factors showed significant results. As already explained in a section above, it can be seen that some company characteristics had an impact on corporate performance, but not sufficiently so to explain the total abnormal return. Consequently, there must be another fact at work such as corporate governance with a significant influence on the positive achievements of a company (Gompers et al., 2003).

5. Conclusion

Sustainability is a complex theme, with no consensus about its performance in the academic world. However, the increasing interest in Socially Responsible Investments has produced a number of researches, which identify potential drivers and new research approaches. By looking at studies with the focus on SRI funds and performance, the picture is mixed. On the one hand, academics believe that environmental, social, and governance performance lead to a competitive advantage and therefore could and should bring new investment opportunities (e.g. Derwall et al., 2005). On the other hand, we see a group of academics, which predict an underperformance for sustainable investments (e.g. Renneboog et al., 2008b) and they often argue with the diversification effect and the efficient market hypothesis. But most of the studies, in the sustainability field, are not able to declare any significance of their results and could be seen as misleading due to the fact that they use funds or indices in their studies. The problem with funds is that the stock-picking ability of the fund manager is captured and not implicitly the effect of sustainability. Due to the use of SRI Indices the academics are able to work around the stock-picking hazard, but they run into the problem of a missing Socially Responsible Investment definition, so they could compare apples with pears. To eliminate the stock picking and the definition issue we focused in our study on sustainability ratings from a worldwide accepted research agency with a clear defined environmental, social, and governance approach.

The research agency “oekom research AG” provided us, on a quarterly basis, with ESG scores for up to 3000 companies from the end of 2005 until December of 2013. These ESG ratings from oekom are constructed out of more then 100 indicators per company. The ESG scores were used to create sustainability portfolios. Thus, all companies were split into sustainable (portfolio B), non-sustainable (portfolio D), and an intermediately portfolio (C). These portfolios were value-weighted and quarterly adjusted for rating changes. Furthermore, to get a better understanding how the sustainability approach reacts in different market situations we split the total sample into a pre-crisis and after-crisis sample. Portfolios were then tested on their performance with a head-to-head approach, Jensen alpha model, and the three-factor model from Fama and French. In all methods the MSCI World Index served as benchmark.

In two out of three samples the results indicated that an outperformance could be generated by investing in the sustainable portfolio B, whereby the significance of the results were mixed. The sustainable portfolio B outperformed in all samples and approaches the other two portfolios, but underperformed the benchmark in the after-crisis sample. Over the total period the sustainable B portfolio was able to generate a return of 40.13%, which is a relative outperformance of 18.23% towards the MSCI World Index. In comparison to that the non-sustainable portfolio D generated a return of 15.51% in the same period, which leads to a relative underperformance of 6.38% towards the MSCI World Index and an underperformance of 24.62% towards the sustainable portfolio B. As nice as these results sound, we only found significant alphas for two portfolios in the Jensen model approach and for one portfolio in the three-factor model over all sample periods. In the Jensen model the alpha from portfolio B was significant in the pre-crisis sample and the alpha from portfolio C was highly significant in the after crisis sample. Furthermore, portfolio C had a significant negative alpha in the three-factor model.

During our study and due to our results a number of further research questions came up. On the one side we found out, that all coefficients of the small-minus-big variable were negative and the question arises, why it is more likely that large firms receive a sustainability ratings compared to small firms. Another question, which came up in our study is, if there is a correlation between the ratings in the different ESG dimensions. Furthermore, are all 100 indicators, used by oekom to identify sustainability of a firm, performance relevant? Respectively, do E, S, and G bring performance or is just one of the dimensions performance relevant? Due to the last research question, which came up during our study, we started a literature review of the environmental, social, and governance literature and their linkage towards performance to find out which indicators are performance-relevant according to academics.

In the environmental dimension the review literature focused on different measurements of environmental performance. It reached from environmental lawsuits and toxic chemical releases up to research agencies with a clear focus on environmental burdens. All studies were quite similar and therefore state that good environmental performance leads to a financial outperformance. In the social dimension we found literature with a focus on a quantitative and others with a focus on a qualitative approach, whereby the qualitative approach, which was mainly based on employee surveys, showed much better results with a high significance.

These studies indicated that the focus should lie on qualitative measurements to generate an outperformance due to social factors and that quantitative approaches are not performance relevant. The corporate governance dimension had a higher literature coverage compared to the other two dimensions and was clear forward, whereby how to define corporate governance was not clear until 2003. Since the "Governance Index" had been developed by Gompers et al. (2003) many academics used this index to compare the level of corporate governance to performance and a significant number of them found significant results. In a nutshell, investments in companies with strong corporate governance are valuable because managers use corporate cash holdings in a more profitable way.

Overall, our study was able to show an outperformance from Socially Responsible Investments, whereby the results did miss out on significance. Therefore, we focused on the E, S, and G dimension to identify which of them is a performance driver and which is an inhibitor. Due to this we were able to identify a number of performance driver for each dimension. For further research the identified indicators could be taken and extracted from research agencies to perform a performance analysis on a large number of companies. Furthermore, the quantitative approach of research agencies could be analysed deeply to find out which of the 100 quantitative indicators are actually performance relevant and how they are correlated to each other. The last research question, which came up during our studies is the fact that sustainable investments perform better in bear markets and underperform in bull markets.

6. Literature

4. S. 2567, 96th Cong., 2d Sess. (1980): ("Protection of Shareholders' Rights Act Of 1980"); H.R. 7010, 96th Cong., 2d Sess. (1980) ("Corporate Democracy Act of 1980").

Aburdene P. (2012): The Global Megatrend of Conscious Capitalism, [online] Available at: www.expomin.cl/marketing/pdf/2012/presentacion_patricia_aburdene.pdf [last access: 12 July 2014]

Akerlof G. (1982): Labour Contracts as Partial Gift Exchanges, The Quarterly Journal of Economics, 97(4), pp. 543-569

Akerlof G. and Yellen J. (1986): The Fair Wage-Effort Hypothesis and Unemployment, The Quarterly Journal of Economics, 105(2); pp. 255-283

Antonasov V. (2005): How much Value can Blockholders Tunnel? Evidence from the Bulgarian Mass Privatization Auctions, Journal of Financial Econometrics, 76(1), pp. 191-234

Barnet M. and Solomon R. (2006): Beyond Dichotomy: The Curvilinear Relationship between Social Responsibility and Financial Performance, Strategy Management Journal, 27(11), pp. 1101-1122

Bauer R., Derwall J., and Otten R. (2007): The Ethical Mutual Fund Performance Debate: New Evidence from Canada, Journal of Business Ethics, 70(2), pp. 111-124

Bauer R., Hann D., and Derwall J. (2009): Employee Relations and Credit Risk, Working Paper presented at the 2009 Academy of Management Meeting, [online] Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1483112 [last access: 7 May 2014]

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

Bauer R., Otten R., and Rat A. (2006): Ethical Investing in Australia: Is there a financial penalty?, Pacific-Basin Finance Journal, 14(1), pp. 33-48

Bebchuk L., Cohen A. and Ferrell A. (2009): What matters in Corporate Governance?, Review of Financial Studies, 22(2), pp. 783-827

Benson K. and Humphrey J. (2007): Socially responsible investment funds: Investor reaction to current and past returns, Journal of Banking & Finance, 32(9), pp. 1850-1859

Brammer S., Brooks C., and Pavelin S. (2006): Corporate social performance and stock return: UK evidence from disaggregate measures, Financial Management, 35(3), pp. 97-116

Brown S. and Goetzmann W. (1995): Performance Persistence, Journal of Finance, 50(2), pp. 679-698

Carhart M. (1997): On Persistence in Mutual Fund Performance, Journal of Finance, 52(1), 57-82

Carroll, A. (1991): The pyramid of Corporate Social Responsibility: Toward the Moral Management of Organizational Stakeholders, Business Horizon, 34(4), pp. 39-48

Core J., Guay W., and Rusticus T. (2006): Does Weak Governance Cause Weak Stock Returns? An Examination of Firm operating Performance and Investors' Expectations, Journal of Finance, 64(2), pp. 655-687

Core J., Holthausen R., and Larcker D. (1999): Corporate governance, chief executive officer compensation, and firm performance, Journal of Financial Economics, 51(3), pp. 371-406

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

Cortez M., Silva F., and Areal N. (2012): Socially Responsible Investing in the Global Market: The Performance of US and European Funds, International Journal of Finance and Econometrics, 17(3), pp. 254-271

Cremers M. and Ferrell A. (2010): Thirty Years of Shareholder Rights and Firm Valuation, Working Paper, Forthcoming in the Journal of Finance, [online] Available at: http://www.fma.org/Hamburg/Papers/Cremers_Ferrell_THIRTY_YEARS_OF_CORPORATE_GOVERNANCE_Nov_09.pdf [last access: 7 May 2014].

Derwall J., Günster N., Bauer R., and Koedijk K. (2005): The Eco-Efficiency Premium Puzzle, Financial Analyst Journal, 61(2), pp. 51-63

Deutsche Bank (2012): Sustainable Investing: Establishing Long-Term Value and Performance, [online] Available at: https://www.dbadvisors.com/content/_media/Sustainable_Investing_2012.pdf [last access: 12 July 2014]

Dittmar A. and Mahrt-Smith J. (2007): Corporate Governance and the Value of Cash Holdings, Journal of Financial Economics, 83(3), pp. 599-634

Dowell G., Hart S., and Yeung B. (2000): Do corporate global environmental standards create or destroy market value?, Management Science, 46(8), pp.1059–1074

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

Edmans A. (2009): Does the stock market fully value intangibles? Employee satisfaction and equity prices, Working Paper, [online] Available at: www.top1000funds.com/attachments/208_DoestheStockMarketFullyValueIntangibles.pdf [last access: 15 July 2014]

Ethical Investment Research Services and UK Social Investment Forum [EIRIS/UKSIF] (2006): The Responsible Investment Approaches of Common Investment Funds, [online] Available at: http://www.eiris.org/files/public%20information%20type%20publications/commoninvestmentfunds06.pdf [last access: 25 October 2014]

Entine, J. (2003): The myth of Social Investing, Organization & Environment, 16 (3), pp. 352-368

Environmental Protection Agency, EPA (2008): Environmentally Preferable Purchasing, [online] Available at: http://www.epa.gov/epp/index.htm [last access: 15 July 2014]

Fama E. and French K. (1993): Common Risk Factors in the Returns on Stocks and Bonds, Journal of Financial Economics, 33(1), pp. 3-56

Faulkeneder M. and Wang R. (2005): Corporate Financial Policy and the Value of Cash, Journal of Finance, 61(4), pp. 1957-1990

Ferson W. and Schadt R (1996): Measuring Fund Strategy and Performance in Changing Economic Conditions, Journal of Finance, 51(2) pp. 425-461

Fischel D. (1982): The Corporate Governance Movement, Vanderbilt Law Review, 35(6), pp. 1259-1292

Fischer, E. and Keber C. (1997): Risikoanalyse internationaler Aktienportefeuilles: Eine empirische Untersuchung, Zeitschrift für Betriebswirtschaft, 67 (3), pp. 333-360

Fisher C. and Lovell A. (2006): Business Ethics and Values, 2nd Edition, Prentice Hall/Harlow, Pearson Education Limited

Friedman M. (1970): The Social Responsibility of Business is to Increase Its Profits, New York Times Magazine 113.IX.1970

Fulmer I., Gerhart B., and Scott K. (2003): Are the 100 Best Better? An Empirical Investigation of the Relation Between Being a "Great Place to Work" and Firm Performance, Personal Psychology, 56(4), pp. 965-993

Fung H-G., Law S., and Yau J. (2010): Socially Responsible Investments: In a Global Environment, Northampton, Edward Elgar Publishing

Galema R., Plantinga A., and Scholtens B. (2008): The stock at stake: Return and risk in socially responsible investment, Journal of Banking and Finance, 32(12), pp. 2646-2654

Global Sustainable Investment Alliance (GSIA) (2013): Global Sustainable Investment Review 2012, [online] Available at: http://gsiareview2012.gsi-alliance.org/pubData/source/Global%20Sustainable%20Investement%20Alliance.pdf [last access: 12 July 2014]

Gompers P., Ishii J., and Metrick A. (2003): Corporate Governance and Equity Prices, The Quarterly Journal of Economics, 118(1), pp. 107-156

Haigh M. and Hazelton J. (2004): Financial Market: A Tool for Social Responsibility?, Journal of Business Ethics, 52(1), pp. 59-71

Hanf J., Iselborn M., and Jungbluth L. (1999): How to deal with the double principal agent problem – a discussion on German wine cooperatives, [online] Available at: http://www.embassyconferences.co.za/media/file_icons/Hanf%20et%20al.pdf [last access: 19 July 2014]

Hart S. and Ahuja G. (1995): Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance, Business Strategy and the Environment, 5, pp. 30-37

Hart S. (1997): Beyond Greening: Strategies for a Sustainable World, Harvard Business Review, 75(1), pp. 66

Herzberg F., Mausner B., and Snyderman B. (1959): The motivation to work, 2nd Edition, New-York: John Wiley & Sons

Hoepner A. and McMillan D. (2009): Research on Responsible Investment: An Influential Literature Analysis Comparing a Rating, Characterisation, Categorisation & Investigation, Working paper series, [online] Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1284 [last access: 12 July 2014]

Hoepner A. and Zeume S. (2009): The dark enemy of responsible mutual funds: does the Vice Fund offer more financial virtue?, Working Paper, [online] Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1485846 [Accessed: 7 May 2014]

Hillier D., Ross S., Westerfield R., Jaffe J., and Jordan B. (2010): Corporate Finance, 1St european Edition, Maidenhead, McGrad-Hill

Hudson J. (2006): The Social Responsibility of the Investment Profession. Research Foundation Publications – CFA Institute, 3, pp. 1-113

Jacobs B., Singhal V., and Subramanian R. (2010): An empirical investigation of environmental performance and the market value of the firm, Journal of Operations Management, 28(5), pp. 430-441

Jensen M. (1968): The Performance of Mutual Funds in the Period 1945-1964, Journal of Finance, 23(2), pp. 389-416

Jensen M. (1986): Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, The American Economic Review, 76(2), pp. 323-329

Kaplan S. and Schweser F. (2012): Corporate Finance, Portfolio Management and Equity Investments, United States, Kaplan Inc.

Kaplan S. and Zingales L. (1997): Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?, The Quarterly Journal of Economics, 112(1), pp. 169-215

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

Kempf A. and Osthoff P. (2007): The Effect of Socially Responsible Investing on Portfolio Performance, European Financial Management, 13(5), pp. 908-922

King A. and Lenox M. (2001): Does it Really Pay to be Green, The Journal of Industrial Ecology, 5(1), pp. 105-116

Klassen R. and McLaughlin C. (1996): The Impact of Environmental Management on Firm Performance, Management Science, 42(8), pp. 1199-1214

Konar S. and Cohen M. (1997): Information as Regulation: The Effect of Community Right to Know Laws on Toxic Emission, Journal of Environmental Economics and Management 32(1), pp. 109-124

Konar S. and Cohen M. (2001): Does the Market Value Environmental Performance, The Review of Economics and Statistics, 83(2), pp. 281-289

Koys D. (2001): The Effects of Employee Satisfaction, Organizational Citizenship Behavior, and Turnover on Organizational Effectiveness: A Unit-Level, Longitudinal Study, Personal Psychology, 54(1), pp. 101-104

Lau R. and May B. (1998): A Win-Win Paradigm for Quality of Work Life and Business Performance, Human Resource Development Quarterly, 9(1), pp. 211-226

Lee D. and Faff R. (2009): In RCM Sustainability White Paper (2011) [online] Available at: http://efnorthamerica.com/documents/events/ccc2008/Mark-Kramer-Keynote/Strategy-Society.PDF [last access: 7 May 2014]

Louche C. and Lydenberg S. (2006): Socially Responsible Investment: Differences between Europe and United States. Vlerick Leuven Gent Working Paper Series 2006/22, pp. 1-37, [online] Available at: http://www.vlerick.be/en/2472-VLK/version/default/part/AttachmentData/data/vlgms-wp-2006-22.pdf [last access: 24 October 2014]

Markowitz H. (1952): Portfolio Selection, Journal of Finance, 7(1), pp. 77-91

McGregor D. (1957): The Human Side of Enterprise, Management Review, 24(1), pp. 41-49

Nelson K. (1994): Finding and Implementing Projects that Reduce Waste, Industrial Ecology and Global Change, 1st Edition, Cambridge, Cambridge University Press

oekom research AG (2014a): Company Presentation on oekom Research Methodology, received by Mr. Boenning, COO/Head of Research, Receive in March 2014

oekom research AG (2014b): Outperformance durch Nutzung des oekom Prime-Standards, oekom Facts & Figures, 1st Edition, [online] Available at: http://www.oekom-research.com/index.php?content=facts-figures [last access: 10 March 2014]

Pearce J. (2003): Social enterprise in Anytown, London, Calouste Gulbenkian Foundation

Pinner W. (2008): ESPA Nachhaltige Investments; ESPA-Vinis Nachhaligkeitsfonds, [online] Available at: http://www.vinis.at/pdf/espa_vinis_folder_D.pdf [last access: 10 June 2014]

Porter M. and van der Linde C. (1995): Green and competitive: ending the stalemate. Harvard Business Review, 73(5), pp. 120–134

Porter M. and Kramer M. (2006): HBR Strategy and Society: The link between competitive advantage and corporate social responsibility, Harvard Business Review, [online]

Available at: http://efnorthamerica.com/documents/events/ccc2008/Mark-Kramer-Keynote/Strategy-Society.PDF [last access: 7 May 2014].

Principles for Responsible Investment (2009): Responsible investment and investment performance, [online] Available at: http://www.unpri.org/viewer/?file=wp-content/uploads/5.Responsibleinvestmentandinvestmentperformance.pdf [last access: 7 May 2014].

Renneboog L. and Zhang C. (2008a): Socially Responsible Investments: Institutional Aspects, Performance, and Investor Behavior. Journal of Banking & Finance, 32(9), pp. 1723-1742

Renneboog L. and Zhang C. (2008b): The Price of Ethics and Stakeholder Governance: The performance of socially responsible mutual funds, Journal of Corporate Finance, 14(3), pp. 302-322

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

Schroeder M. (2007): Is there a Difference? The Performance Characteristics of SRI Equity Indices, Journal of Business Finance & Accounting, 32(1/2), pp. 331-348

Schueth M. (2007): Socially Responsible Investing in the United States, Journal of Business Ethics, 43(3), pp. 189-194

Sethi S. (2005): Investing in Socially Responsible Companies is a Must for Public Pension Funds – Because There is no Better Alternative, Journal of Business Ethics, 56(2), pp. 99-129

Shapiro C. and Stiglitz E. (1984): Equilibrium Unemployment as a Worker Discipline Device, American Economic Association, 74(3), pp. 422-444

Sharpe W. (1964): Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, Journal of Finance, 19(3), pp. 425-442

Sparkes R. (2001): Ethical Investment: Whose Ethics, which Investment?, Business Ethics: A European Review, 10(3), 194-205

Spicer B. (1978): Investors, Corporate Social Performance, and Informational Disclosure: An Empirical Study, Accounting Review, 53(1), pp. 94-103

Sroufe R. (2003): Effects of environmental management systems on environmental management practices and operations, Production and Operations Management, 12(3), pp. 416–431

Statman M. (2000): Socially Responsible Mutual Funds, Financial Analysts Journal, 56(3), pp. 30-39

Statman M. (2006): Socially Responsible Indexes: Composition, performance and tracking error, The Journal of Portfolio Management, 32(3), pp. 100-109

Stenstöm H. and Thorell J. (2007): Evaluating the performance of socially responsible investment funds: A holding data analysis, Journal of Business Ethics, 87(4), pp. 573-588

Trevino L. and Nelson K. (2007): Managing Business Ethics, 3rd Edition, John Wiley & Sons Inc., Westwood/Hoboken

USSIF (2010): Report on Socially Responsible Investing Trends in the United States, [online] Available at: http://www.ussif.org/files/publications/12_trends_exec_summary.pdf [last access: 12 July 2014]

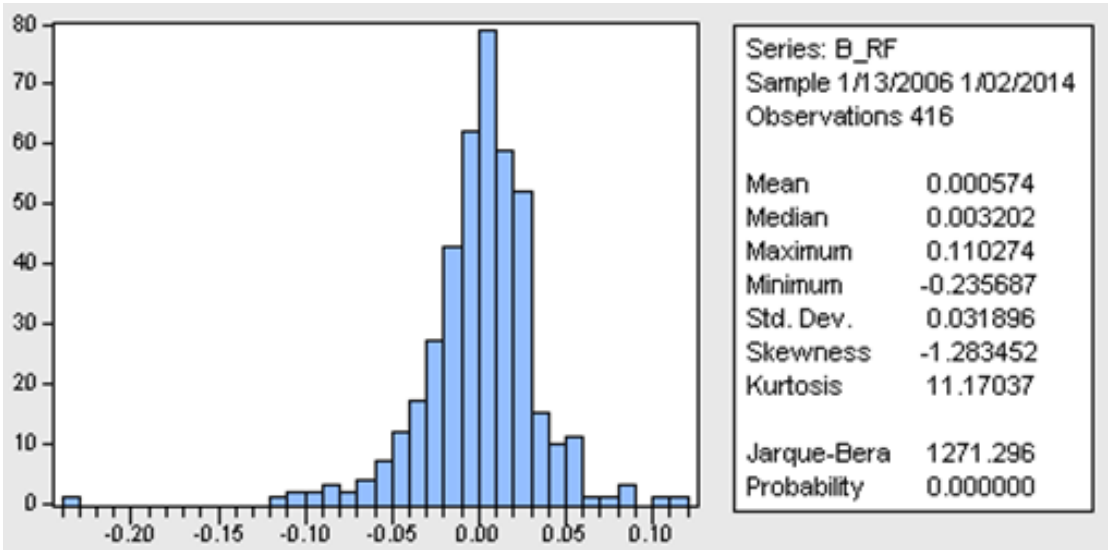
Vogan C. (1996): Pollution Abatement and Control Expenditures 1972-94, Survey of Current Business 76(9), pp. 48-67

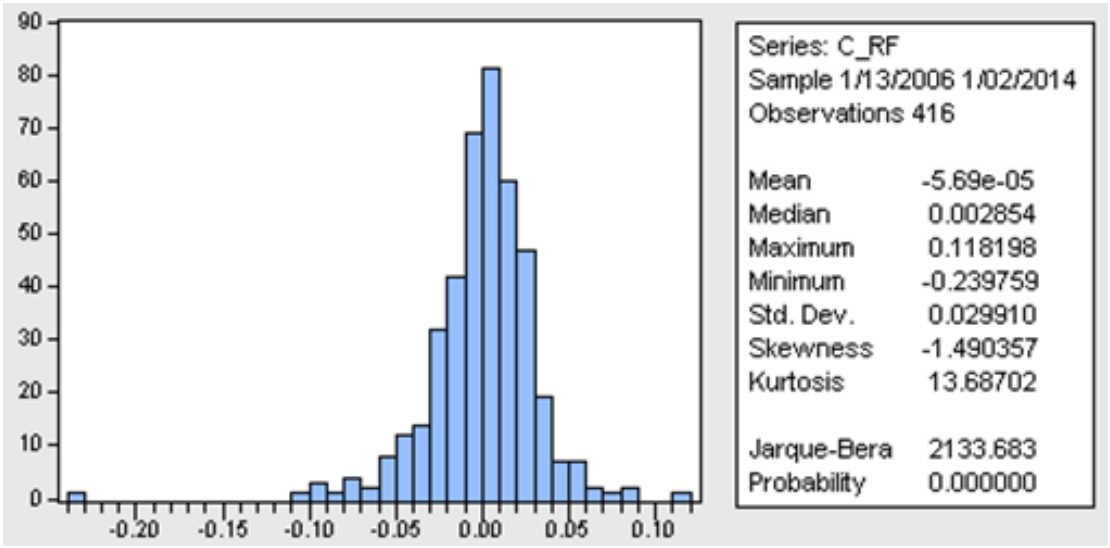
Waddock S. (2003): Unfolding Stakeholder Thinking, Greenleaf Publishing in association with GSE Research, [online] Available at: http://www.ingentaconnect.com/content/glbj/ust/2002/00000001/00000121/art00001 [last access: 12 July 2014]

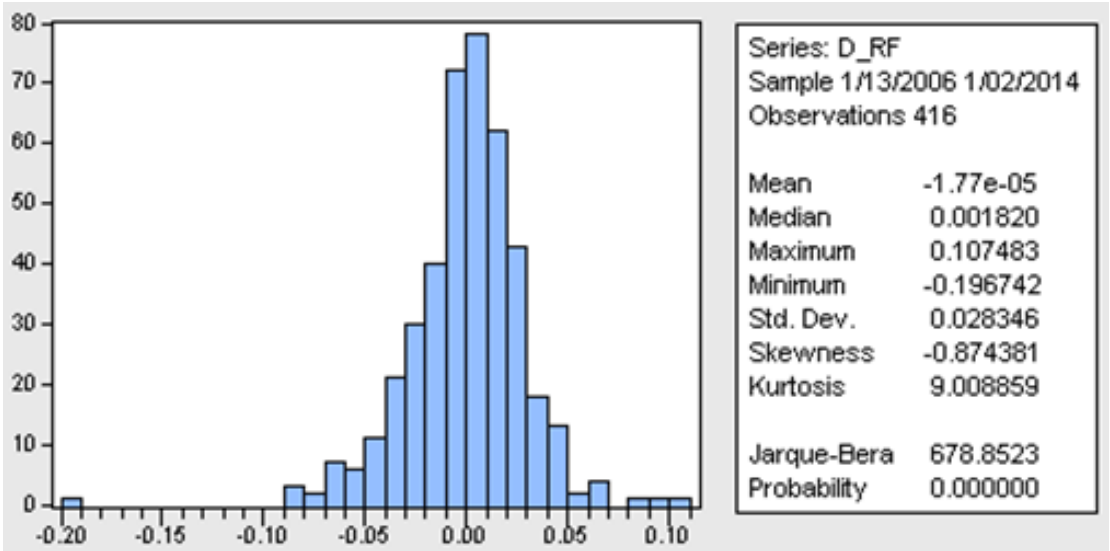
Yermack D. (2006): Flights of Fancy: Corporate Jets, CEO Perquisites, and Inferior Shareholder Return, Journal of Financial Econometrics, 80(1), pp. 211-242

7. Appendix

Appendix I.I – Total Sample Statistics







Appendix I.II – Total Sample CAPM

Dependent Variable: B_RF
Method: Least Squares
Date: 09/03/14 Time: 20:24
Sample: 1/13/2006 12/27/2013
Included observations: 416

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000386	0.000395	0.975828	0.3297
MARKT_3MEUR	1.079977	0.013846	78.00190	0.0000

R-squared	0.936291	Mean dependent var	0.000574
Adjusted R-squared	0.936137	S.D. dependent var	0.031896
S.E. of regression	0.008060	Akaike info criterion	-6.798915
Sum squared resid	0.026897	Schwarz criterion	-6.779537
Log likelihood	1416.174	Hannan-Quinn criter.	-6.791253
F-statistic	6084.297	Durbin-Watson stat	2.144130
Prob(F-statistic)	0.000000		

Dependent Variable: C_RF
Method: Least Squares
Date: 09/03/14 Time: 20:25
Sample: 1/13/2006 12/27/2013
Included observations: 416

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000238	0.000203	-1.171393	0.2421
MARKT_3MEUR	1.036574	0.007113	145.7223	0.0000

R-squared	0.980877	Mean dependent var	-5.69E-05
Adjusted R-squared	0.980831	S.D. dependent var	0.029910
S.E. of regression	0.004141	Akaike info criterion	-8.130892
Sum squared resid	0.007100	Schwarz criterion	-8.111514
Log likelihood	1693.226	Hannan-Quinn criter.	-8.123230
F-statistic	21234.98	Durbin-Watson stat	2.136042
Prob(F-statistic)	0.000000		

Dependent Variable: D_RF
Method: Least Squares
Date: 09/03/14 Time: 20:32
Sample: 1/13/2006 12/27/2013
Included observations: 416

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000178	0.000526	-0.338692	0.7350
MARKT_3MEUR	0.918444	0.018412	49.88236	0.0000

R-squared	0.857352	Mean dependent var	-1.77E-05
Adjusted R-squared	0.857007	S.D. dependent var	0.028346
S.E. of regression	0.010719	Akaike info criterion	-6.228813
Sum squared resid	0.047567	Schwarz criterion	-6.209435
Log likelihood	1297.593	Hannan-Quinn criter.	-6.221151
F-statistic	2488.250	Durbin-Watson stat	2.276639
Prob(F-statistic)	0.000000		

Appendix I.III – Total Sample Fama & French

Dependent Variable: B_RF
Method: Least Squares
Date: 09/03/14 Time: 20:33
Sample: 1/13/2006 12/27/2013
Included observations: 416

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000274	0.000366	0.747352	0.4553
MARKT_3MEUR	1.126858	0.014016	80.40042	0.0000
SMB	-0.259144	0.052516	-4.934546	0.0000
HML	-0.397804	0.048232	-8.247752	0.0000

R-squared	0.945754	Mean dependent var	0.000574
Adjusted R-squared	0.945359	S.D. dependent var	0.031896
S.E. of regression	0.007456	Akaike info criterion	-6.950091
Sum squared resid	0.022902	Schwarz criterion	-6.911334
Log likelihood	1449.619	Hannan-Quinn criter.	-6.934767
F-statistic	2394.330	Durbin-Watson stat	2.232750
Prob(F-statistic)	0.000000		

Dependent Variable: C_RF
Method: Least Squares
Date: 09/03/14 Time: 20:35
Sample: 1/13/2006 12/27/2013
Included observations: 416

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000126	0.000187	-0.672751	0.5015
MARKT_3MEUR	1.034684	0.007160	144.5176	0.0000
SMB	-0.140085	0.026827	-5.221830	0.0000
HML	0.112901	0.024638	4.582345	0.0000

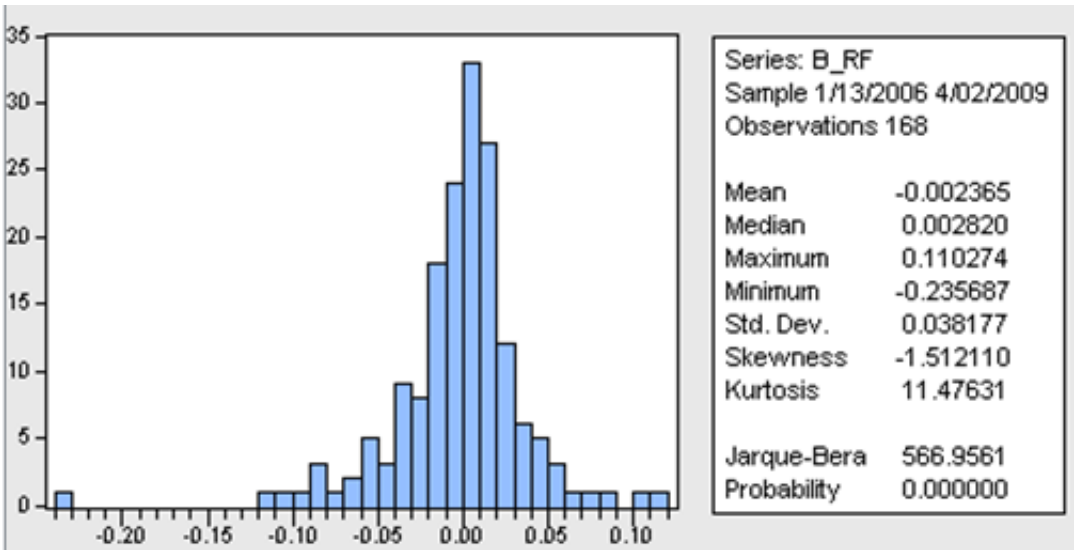
R-squared	0.983903	Mean dependent var	-5.69E-05
Adjusted R-squared	0.983785	S.D. dependent var	0.029910
S.E. of regression	0.003809	Akaike info criterion	-8.293529
Sum squared resid	0.005976	Schwarz criterion	-8.254772
Log likelihood	1729.054	Hannan-Quinn criter.	-8.278205
F-statistic	8394.086	Durbin-Watson stat	2.166030
Prob(F-statistic)	0.000000		

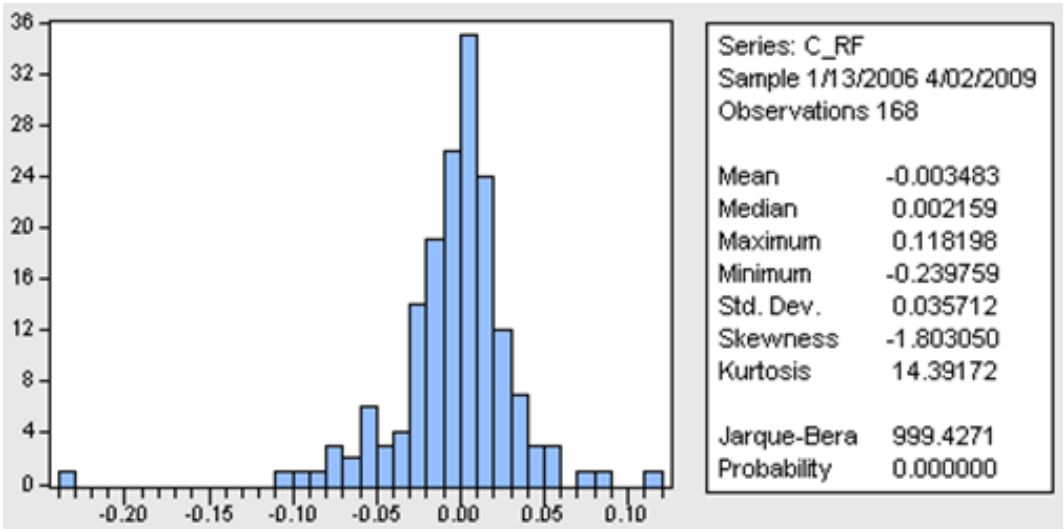
Dependent Variable: D_RF
Method: Least Squares
Date: 09/03/14 Time: 20:35
Sample: 1/13/2006 12/27/2013
Included observations: 416

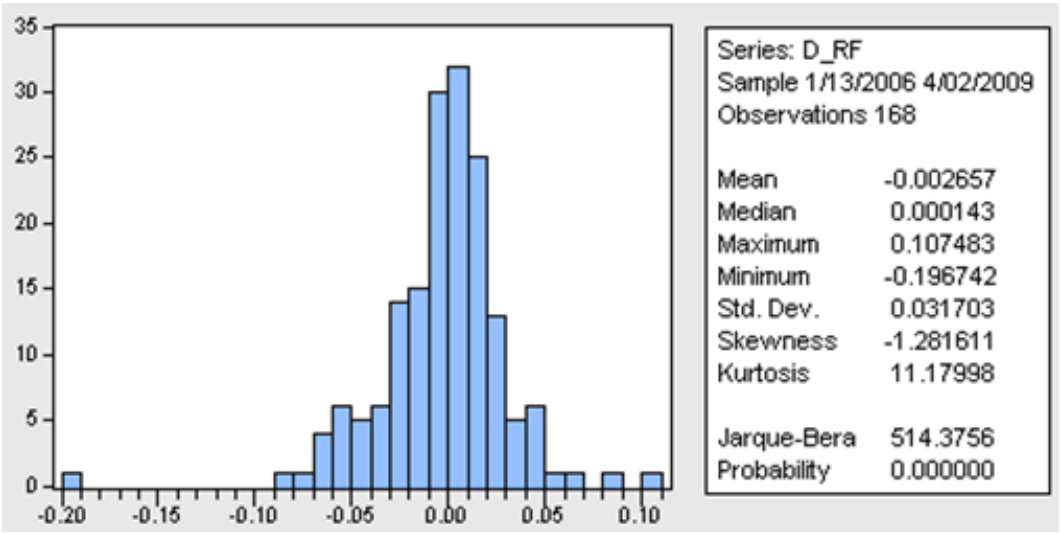
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000112	0.000526	-0.213392	0.8311
MARKT_3MEUR	0.914523	0.020120	45.45275	0.0000
SMB	-0.057372	0.075391	-0.761000	0.4471
HML	0.084129	0.069240	1.215031	0.2250

R-squared	0.858457	Mean dependent var	-1.77E-05
Adjusted R-squared	0.857426	S.D. dependent var	0.028346
S.E. of regression	0.010703	Akaike info criterion	-6.226971
Sum squared resid	0.047199	Schwarz criterion	-6.188214
Log likelihood	1299.210	Hannan-Quinn criter.	-6.211646
F-statistic	832.9220	Durbin-Watson stat	2.303365
Prob(F-statistic)	0.000000		

Appendix II.I – Pre Crisis Statistics







Appendix II.II – Pre Crisis CAPM

Dependent Variable: B_RF
Method: Least Squares
Date: 09/03/14 Time: 20:45
Sample: 1/13/2006 3/27/2009
Included observations: 168

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001361	0.000760	1.790920	0.0751
MARKT_3MEUR	1.067989	0.021940	48.67687	0.0000

R-squared	0.934528	Mean dependent var	-0.002365
Adjusted R-squared	0.934134	S.D. dependent var	0.038177
S.E. of regression	0.009798	Akaike info criterion	-6.401455
Sum squared resid	0.015936	Schwarz criterion	-6.364265
Log likelihood	539.7222	Hannan-Quinn criter.	-6.386361
F-statistic	2369.437	Durbin-Watson stat	2.262987
Prob(F-statistic)	0.000000		

Dependent Variable: C_RF
Method: Least Squares
Date: 09/03/14 Time: 20:47
Sample: 1/13/2006 3/27/2009
Included observations: 168

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.87E-05	0.000377	0.235235	0.8143
MARKT_3MEUR	1.023861	0.010884	94.07038	0.0000

R-squared	0.981587	Mean dependent var	-0.003483
Adjusted R-squared	0.981476	S.D. dependent var	0.035712
S.E. of regression	0.004860	Akaike info criterion	-7.803527
Sum squared resid	0.003922	Schwarz criterion	-7.766337
Log likelihood	657.4963	Hannan-Quinn criter.	-7.788433
F-statistic	8849.236	Durbin-Watson stat	2.190000
Prob(F-statistic)	0.000000		

Dependent Variable: D_RF
Method: Least Squares
Date: 09/03/14 Time: 20:47
Sample: 1/13/2006 3/27/2009
Included observations: 168

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000354	0.000836	0.423579	0.6724
MARKT_3MEUR	0.863067	0.024145	35.74458	0.0000

R-squared	0.885016	Mean dependent var	-0.002657
Adjusted R-squared	0.884323	S.D. dependent var	0.031703
S.E. of regression	0.010783	Akaike info criterion	-6.209924
Sum squared resid	0.019300	Schwarz criterion	-6.172734
Log likelihood	523.6336	Hannan-Quinn criter.	-6.194830
F-statistic	1277.675	Durbin-Watson stat	2.566350
Prob(F-statistic)	0.000000		

Appendix II.III – Pre Crisis Fama & French

Dependent Variable: B_RF
Method: Least Squares
Date: 09/03/14 Time: 20:48
Sample: 1/13/2006 3/27/2009
Included observations: 168

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000920	0.000649	1.418094	0.1581
MARKT_3MEUR	1.127539	0.020046	56.24802	0.0000
SMB	-0.366195	0.081543	-4.490833	0.0000
HML	-0.579914	0.072113	-8.041790	0.0000

R-squared	0.953258	Mean dependent var	-0.002365
Adjusted R-squared	0.952403	S.D. dependent var	0.038177
S.E. of regression	0.008329	Akaike info criterion	-6.714619
Sum squared resid	0.011377	Schwarz criterion	-6.640239
Log likelihood	568.0280	Hannan-Quinn criter.	-6.684432
F-statistic	1114.868	Durbin-Watson stat	2.379575
Prob(F-statistic)	0.000000		

Dependent Variable: C_RF
Method: Least Squares
Date: 09/03/14 Time: 20:49
Sample: 1/13/2006 3/27/2009
Included observations: 168

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.79E-05	0.000345	0.138692	0.8899
MARKT_3MEUR	1.015790	0.010670	95.19832	0.0000
SMB	-0.155360	0.043405	-3.579347	0.0005
HML	0.102745	0.038385	2.676699	0.0082

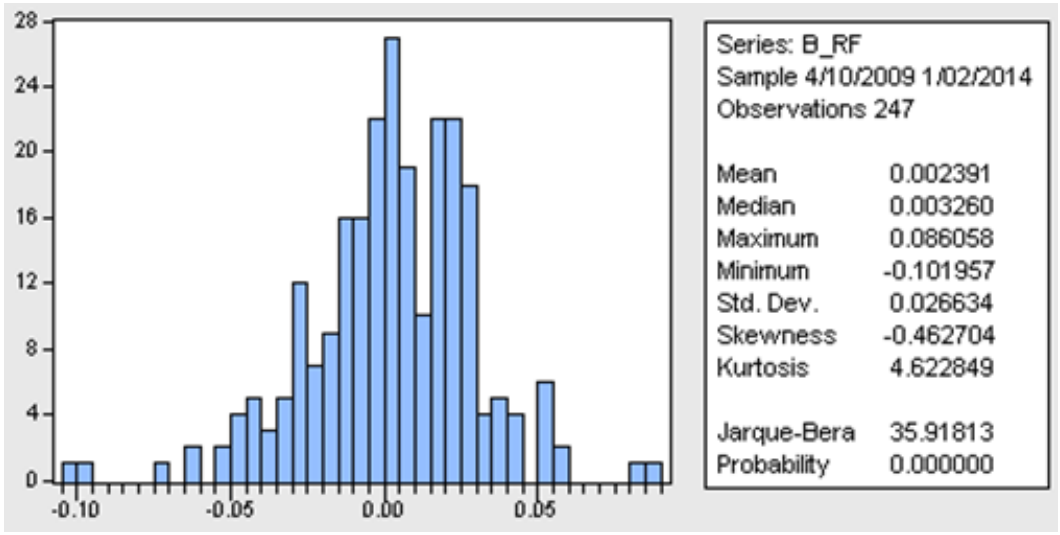
R-squared	0.984864	Mean dependent var	-0.003483
Adjusted R-squared	0.984588	S.D. dependent var	0.035712
S.E. of regression	0.004433	Akaike info criterion	-7.975742
Sum squared resid	0.003224	Schwarz criterion	-7.901362
Log likelihood	673.9624	Hannan-Quinn criter.	-7.945555
F-statistic	3557.138	Durbin-Watson stat	2.165956
Prob(F-statistic)	0.000000		

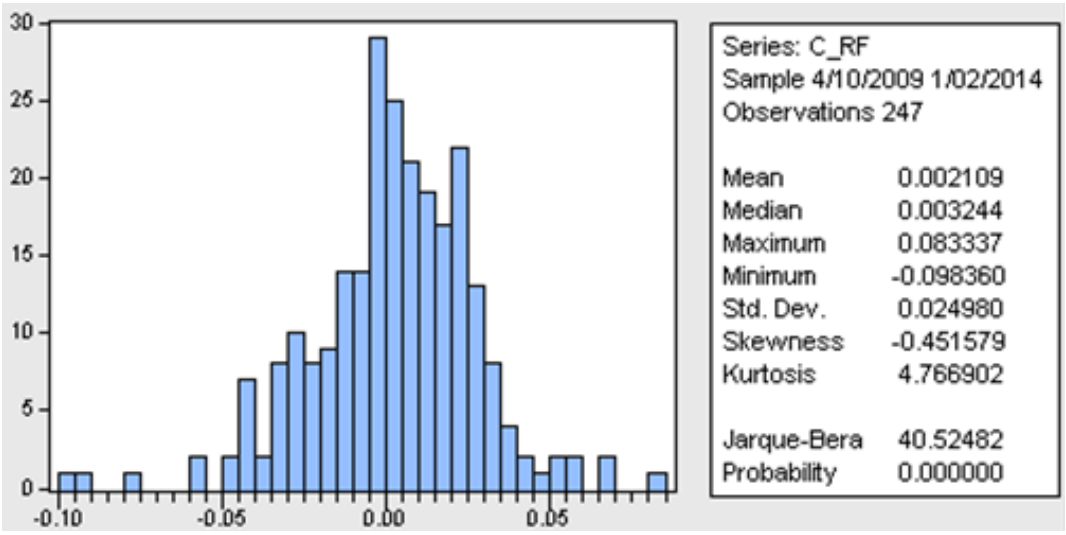
Dependent Variable: D_RF
Method: Least Squares
Date: 09/03/14 Time: 20:49
Sample: 1/13/2006 3/27/2009
Included observations: 168

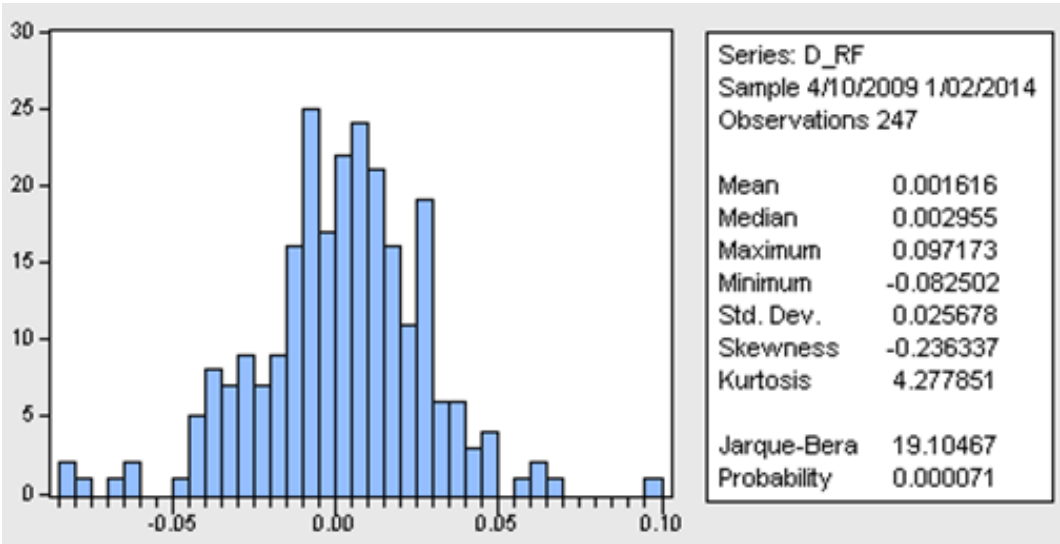
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000200	0.000836	0.239159	0.8113
MARKT_3MEUR	0.876222	0.025817	33.93912	0.0000
SMB	-0.197211	0.105021	-1.877836	0.0622
HML	-0.114401	0.092875	-1.231765	0.2198

R-squared	0.887568	Mean dependent var	-0.002657
Adjusted R-squared	0.885511	S.D. dependent var	0.031703
S.E. of regression	0.010727	Akaike info criterion	-6.208559
Sum squared resid	0.018872	Schwarz criterion	-6.134179
Log likelihood	525.5190	Hannan-Quinn criter.	-6.178372
F-statistic	431.5520	Durbin-Watson stat	2.485782
Prob(F-statistic)	0.000000		

Appendix III.I – After Crisis Statistics







Appendix III.II – After Crisis CAPM

Dependent Variable: B_RF
Method: Least Squares
Date: 09/03/14 Time: 21:02
Sample: 4/10/2009 12/27/2013
Included observations: 247

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000393	0.000419	-0.936371	0.3500
MARKT_3MEUR	1.103749	0.017854	61.82198	0.0000

R-squared	0.939758	Mean dependent var	0.002391
Adjusted R-squared	0.939513	S.D. dependent var	0.026634
S.E. of regression	0.006550	Akaike info criterion	-7.210502
Sum squared resid	0.010513	Schwarz criterion	-7.182085
Log likelihood	892.4969	Hannan-Quinn criter.	-7.199061
F-statistic	3821.957	Durbin-Watson stat	1.922734
Prob(F-statistic)	0.000000		

Dependent Variable: C_RF
Method: Least Squares
Date: 09/03/14 Time: 21:03
Sample: 4/10/2009 12/27/2013
Included observations: 247

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000558	0.000225	-2.476978	0.0139
MARKT_3MEUR	1.057245	0.009590	110.2501	0.0000

R-squared	0.980242	Mean dependent var	0.002109
Adjusted R-squared	0.980161	S.D. dependent var	0.024980
S.E. of regression	0.003518	Akaike info criterion	-8.453577
Sum squared resid	0.003033	Schwarz criterion	-8.425161
Log likelihood	1046.017	Hannan-Quinn criter.	-8.442137
F-statistic	12155.07	Durbin-Watson stat	2.087334
Prob(F-statistic)	0.000000		

Dependent Variable: D_RF
Method: Least Squares
Date: 09/03/14 Time: 21:03
Sample: 4/10/2009 12/27/2013
Included observations: 247

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000916	0.000665	-1.377834	0.1695
MARKT_3MEUR	1.004170	0.028325	35.45219	0.0000

R-squared	0.836869	Mean dependent var	0.001616
Adjusted R-squared	0.836203	S.D. dependent var	0.025678
S.E. of regression	0.010392	Akaike info criterion	-6.287455
Sum squared resid	0.026460	Schwarz criterion	-6.259039
Log likelihood	778.5007	Hannan-Quinn criter.	-6.276015
F-statistic	1256.858	Durbin-Watson stat	2.097639
Prob(F-statistic)	0.000000		

Appendix III.III – After Crisis Fama & French

Dependent Variable: B_RF
Method: Least Squares
Date: 09/03/14 Time: 21:04
Sample: 4/10/2009 12/27/2013
Included observations: 247

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000347	0.000413	-0.839348	0.4021
MARKT_3MEUR	1.136886	0.019817	57.37037	0.0000
SMB	-0.206018	0.068961	-2.987468	0.0031
HML	-0.178899	0.063961	-2.797019	0.0056

R-squared	0.942802	Mean dependent var	0.002391
Adjusted R-squared	0.942096	S.D. dependent var	0.026634
S.E. of regression	0.006409	Akaike info criterion	-7.246154
Sum squared resid	0.009981	Schwarz criterion	-7.189322
Log likelihood	898.9001	Hannan-Quinn criter.	-7.223273
F-statistic	1335.138	Durbin-Watson stat	2.023672
Prob(F-statistic)	0.000000		

Dependent Variable: C_RF
Method: Least Squares
Date: 09/03/14 Time: 21:05
Sample: 4/10/2009 12/27/2013
Included observations: 247

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000373	0.000207	-1.802375	0.0727
MARKT_3MEUR	1.067495	0.009917	107.6454	0.0000
SMB	-0.164691	0.034510	-4.772288	0.0000
HML	0.119526	0.032008	3.734294	0.0002

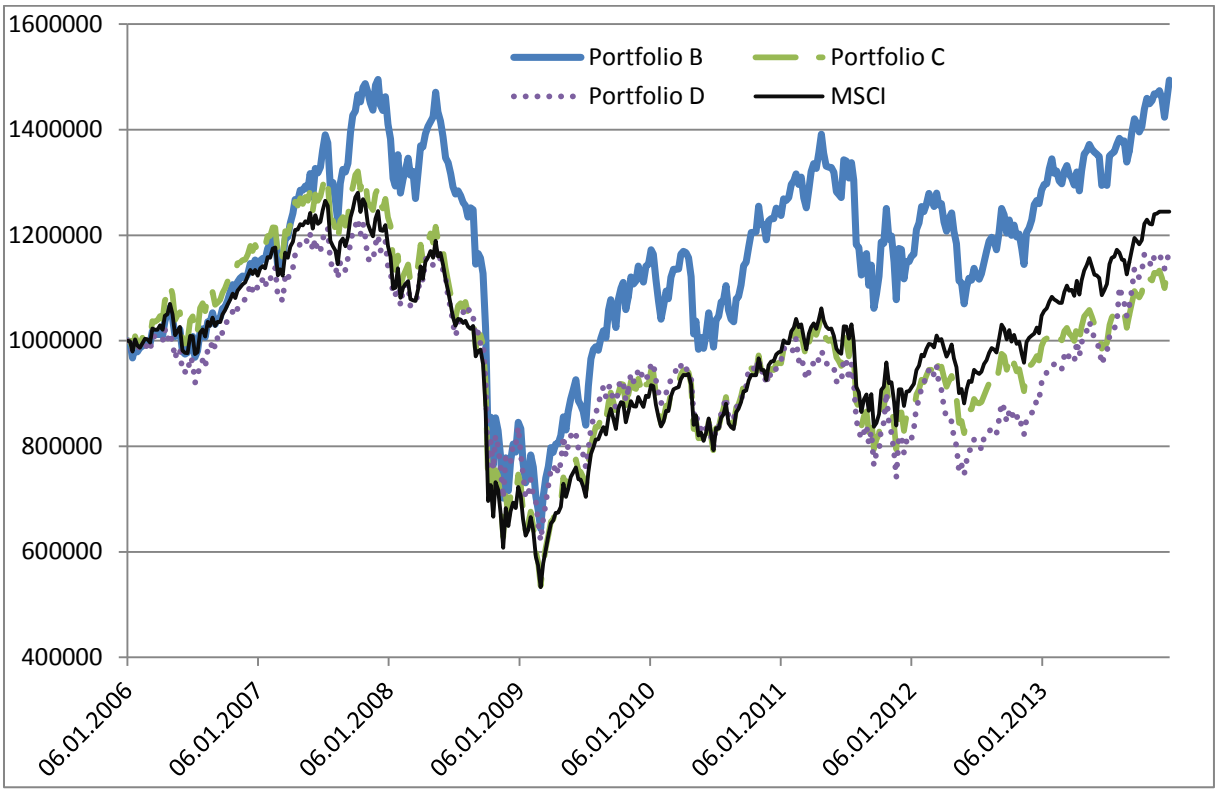
R-squared	0.983716	Mean dependent var	0.002109
Adjusted R-squared	0.983515	S.D. dependent var	0.024980
S.E. of regression	0.003207	Akaike info criterion	-8.630740
Sum squared resid	0.002500	Schwarz criterion	-8.573908
Log likelihood	1069.896	Hannan-Quinn criter.	-8.607859
F-statistic	4893.130	Durbin-Watson stat	2.152779
Prob(F-statistic)	0.000000		

Dependent Variable: D_RF
Method: Least Squares
Date: 09/03/14 Time: 21:05
Sample: 4/10/2009 12/27/2013
Included observations: 247

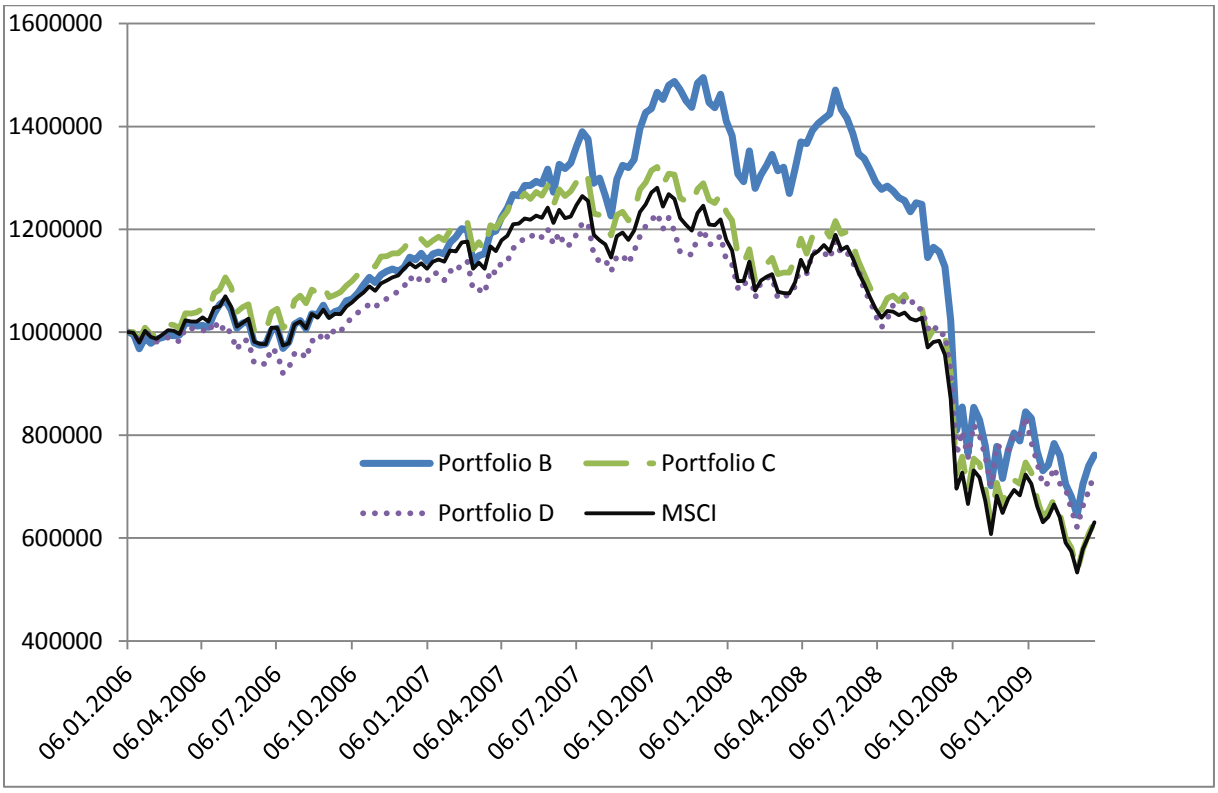
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000701	0.000656	-1.068095	0.2865
MARKT_3MEUR	0.987546	0.031452	31.39887	0.0000
SMB	-0.037603	0.109450	-0.343562	0.7315
HML	0.333874	0.101514	3.288938	0.0012

R-squared	0.844985	Mean dependent var	0.001616
Adjusted R-squared	0.843071	S.D. dependent var	0.025678
S.E. of regression	0.010172	Akaike info criterion	-6.322292
Sum squared resid	0.025143	Schwarz criterion	-6.265460
Log likelihood	784.8030	Hannan-Quinn criter.	-6.299411
F-statistic	441.5288	Durbin-Watson stat	2.076859
Prob(F-statistic)	0.000000		

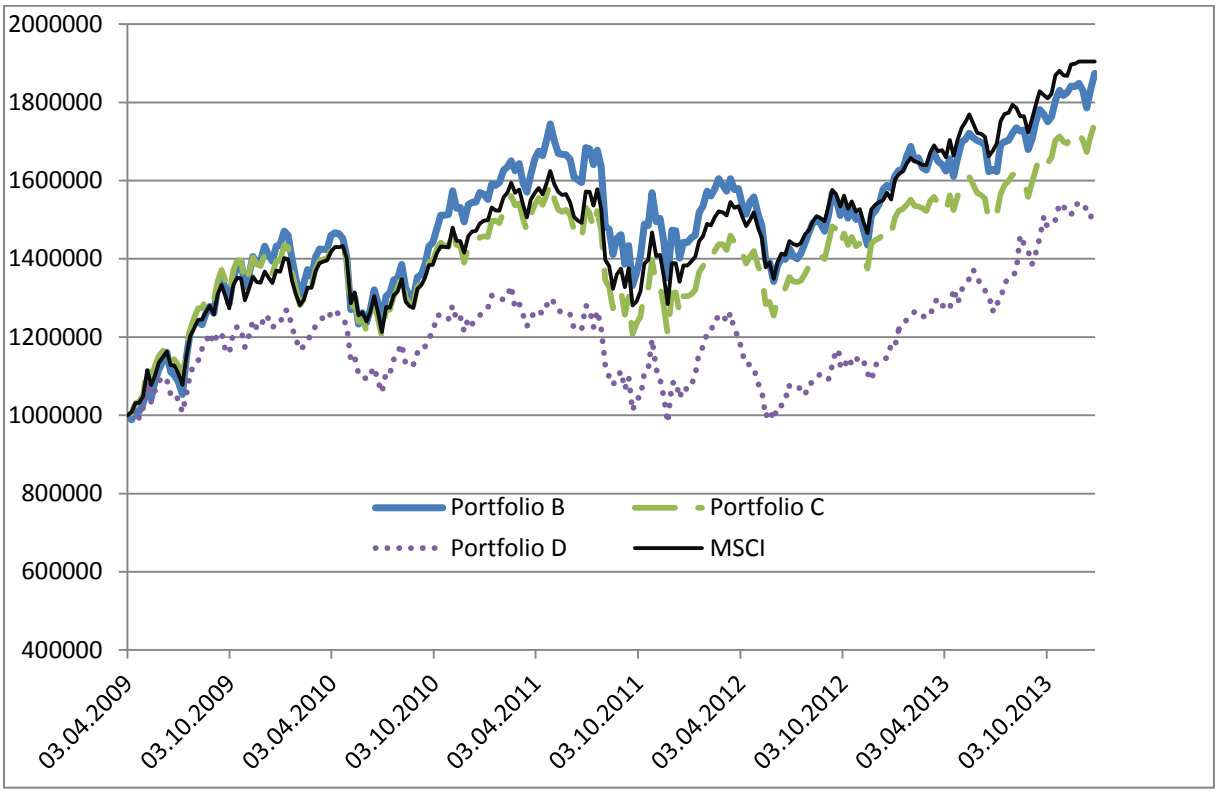
Appendix IV.I – Total Sample Head-to-Head Chart



Appendix IV.II – Pre Crisis Head-to-Head Chart



Appendix IV.III – After Crisis Head-to-Head Chart



Abstract (Deutsch)

In der Finanzanalyse verändert sich derzeit die Nachhaltigkeitsanalyse vom Nischenprodukt zu einem Kernprodukt und generiert hierdurch immer mehr Aufmerksamkeit. In den letzten Jahren ist die Anzahl von Nachhaltigkeitsstudien im Finanzsektor exponentiell angestiegen. Trotz der steigenden Anzahl an Studien konzentrieren sich die meisten Autoren auf nachhaltige Fonds und nachhaltige Indizes, was eine Vergleichbarkeit erschwert. Des Weiteren hat der Markt noch kein Konsensus gefunden ob Nachhaltigkeit eine signifikante Überrendite ermöglicht oder nicht. Diese Arbeit hingegen basiert auf Nachhaltigkeitsdaten eines anerkannten Researchanbieters, der Unternehmen, anhand von über 100 Indikatoren auf Ihre Nachhaltigkeit prüft. Basierend auf ihrer Nachhaltigkeit haben wir die Unternehmen in Portfolios aufgeteilt und anschließend eine Performanceanalyse mit einer Head-to-Head-Analyse, Jensen Model und dem Fama und French Model durchgeführt. Die Analyse basiert auf Aktienrenditen zwischen 2006 und 2014. In unserer Arbeit ist eine deutliche Überrendite durch die nachhaltigsten Unternehmen ersichtlich, wobei nicht von signifikanten Resultaten gesprochen werden kann.

Um herauszufinden woher diese Überrendite stammt wurde eine Vielzahl von Studien zu Umwelt-, sozialen und Governance-Faktoren zusammengefasst. In allen E, S und G Bereichen gibt es Argumente für eine Überrendite. Diese reichen von Innovationsvorteilen zu motivierten Mitarbeitern bis hin zur Überrendite durch einen unabhängigen Aufsichtsrat.

Abstract (English)

Sustainability is currently changing from a niche component to a main component in the financial analysis. Due to this change, sustainability receives more and more attention from academics. Hence, the number of sustainability studies in the financial sector has grown exponentially in recent years. Despite the increasing number of studies, most authors focused on sustainable funds and sustainable indices, which makes a comparison difficult. Furthermore, the market has not found a consensus whether sustainability enables a significant excess return or not. Thus, this study concentrates on a different approach to measure sustainability. We take the sustainability rating from a trusted sustainability research provider and constructed portfolios for the different sustainability ratings. Subsequently, we conducted a performance analysis by analyzing the equity returns with a head-to-head analysis, the Jensen model and the Fama and French model. We found out that the sustainable portfolio outperformed the market, but the results were not significant.

To find out due to which factors the outperformance can be generated, a significant number of environmental, social and governance studies are summarized at the end of the study. We find arguments in favor of outperformance possibilities for all E, S and G criteria. They range from innovation benefits and motivated workers up to an independent board and these might be drivers for a financial outperformance.

Richard Boulanger
Kurzbauergasse 5/6
1020 Wien, Österreich
Mobile : +43 676 755 655 0
r.boulanger89@gmail.com

Persönliche Daten

Geburtsort	Berlin-Wilmersdorf, 13.05.1989

Ausbildung

2012 -	**Universität Wien**, Österreich Studiengang → Betriebswirtschaftslehre MSc., Kernfachkombination: Finanzmärkte und Corperate Finance
2012 - 2014	**Institut für strategische Kapitalmarktforschung Wien**, Österreich Portfolio Management Programm, Verwaltung eines Studentenfonds (1,3 Mio. €, Verfolgung einer Carry-Strategie, alle Anlageklassen)
2009 - 2012	**Universität Wien**, Österreich Abschluss → Internationale Betriebswirtschaftslehre BSc., Kernfachkombination: Public Utility Management, International Financial Markets
2005 - 2008	**Martin-Niemöller-Schule**, Wiesbaden, Deutschland Abschluss → Abitur
2003 - 2005	**École secondaire Cardial Carter** (Secondary), Aurora, Ontario, Canada

Berufserfahrung

02/2014 - jetzt	**Erste Asset Management GmbH**, Wien, Österreich Analyst, Responsible Investments; Neuaufbau der Abteilung, Research, Voting, Engagement, Strategieentwicklung
09/2011 - 02/2014	**Universität Wien**, 1010 Wien, Österreich Studienassistent, Institut für Finanzwirtschaft (*10 Std. Woche)* Fächer: Quantitative Methoden der BWL und Principles of Finance
11/2012 – 04/2013	**tecnet Equity**, Wien, Österreich Analyst, Venture Capital Fonds; Nach Praktikumsende weitere projektbezogene Mitarbeit
02/2011 - 07/2012	**Bank Austria (UniCredit)**, 1090 Wien, Österreich Werksstudent; Kreditierung und Organisation der Archive (*12 Std. Woche)*
01/2009 - 10/2009	**Zivildienst,** Wiesbaden-Biebrich, Deutschland **WfB**, Werkstatt für Behinderte; Unterstützung der Teams und Auftragsakquise

Praktika

09/2013 – 02/2014	**Erste Asset Management GmbH**, Wien, Österreich *(25Std. Woche)* Responsible Investments; Neuaufbau der Abteilung, Research, Voting, Engagement, Strategieentwicklung
08/2012 – 10/2012	**Keller & Coll.**, Frankfurt, Deutschland M&A Beratung; Bearbeitung eines eigenen Projekts (Targetlisten, Bewertung, Markt-, Konkurrenz-, Produktanalysen sowie Erstellung des Info-Memorandums)
07/2012 – 08/2012	**tecnet Equity**, Wien, Österreich Venture Capital Fonds; Unterstützung bei bestehenden und zukünftigen Projekten (Marktanalysen, potenzielle Käufer, Szenarioanalysen)

07/2010 – 09/2010	**CIAT-Group**, Culoz, Frankreich Strategische Unternehmensplanung; Entwicklung neuer potenzieller Märkte anhand von Markt- und Konkurrenzanalysen, Entwicklung eines Makroindikatoren-Baskets
12/2006 – 06/2007	**IHK**, Workshop "Managementpraxis", Wiesbaden, Deutschland Seminare mit Personen aus dem Management um Unternehmensstrategien kennen zu lernen und zu verstehen

Sprachen

Deutsch	Muttersprache
Französisch	Muttersprache
Englisch	verhandlungssicher

Hobbies

Reisen	Städtetouren
Sport	Golf, Handball, Motorradsport, Beachvolleyball

Sonstiges

06/2015	Prüfungsteilnehmer CFA Level 2
Seit 10/2014	Externer Lektor, Universität Wien; Repetitorium für Quantitative Methoden der BWL
05/2014	Zertifizierter EFFAS ESG-Analyst
12/2013	CFA Level 1
05/2013	Mitglied im e-fellows.net Netzwerk (Stipendium)
02/2012	Mitglied in der Class of Excellence

Wien, 2015