Master Thesis

Do sustainable firms induce sustainable outperformance?

M. van de Kerkhof

August 2014

M. van de Kerkhof

542508

Scriptiebegeleider:

Prof. Dr. L.D.R. Renneboog

Master Finance

Departement of Finance

Tilburg University

Helmond, August 2014

Management Summary

Sustainable and responsible investing, also known as social responsible investing (SRI) expanded at a higher pace than all other investment assets from 2007 to 2010. This raises the question on how SRI is best implemented. This study analyzes different screening policies based on KLD database. The sample is restricted to the S&P 500 stocks for the period of 1991 – 2012. Instead of using models such as CAPM or the Carhart 4-factor model, linear regressions are ran in order to determine whether such relations exists between screening policies and returns. The main results are inconclusive. Sin stocks seem to have a higher coefficient on stock returns than non controversial stock. When regressing all three screens on return, the environment screen is not statistically significant it has a small negative coefficient. The coefficient for social screens is also negative, but it is statistically significant. Corporate governance screens are highly statistical significant and have a large positive coefficient. When performing pairs of two screens at once, omitting one screen, coefficients of environment and social screens rise when omitting the governance screen. Similar results appear when omitting environment and social screens. The industry test shows that no statistical significance was found, but in general, industries are positively related to the environment and negatively related to governance screens. Finally a time test is done for sub-periods 1991 – 2002 and 2003 – 2012. Including sin stocks shows a positive relation of environmental and social screens on return, but excluding sin stocks shows a positive relation of environmental and social screens on return.

Foreword

In front of you lays my thesis to round of my Master in Finance at Tilburg University. This thesis attempts to find linear relationships between screening policies and stock returns.

In 2009 I started my master in Financial Management at Tilburg University. Taking various courses on investment analysis and investor behaviour has strongly influenced the choice of my carreer and this thesis. My strong interest for social responsible investing has led me to write my final thesis on this topic.

A growing interest of social responsible investing has not been unnoticed by the financial industry. Many asset managers and financial institutions have created mutual funds in order to satisfy customer demands.

I should end by thanking my thesis supervisor, Prof. Dr. L.D.R. Renneboog for his valuable advice and support and keeping faith in me, even when I tended to lose faith myself. Furthermore I would like to thank everybody who has supported me during my study at Tilburg University. I would like to specially thank my wife Mireille for her support and standing by me during difficult times. Finally I would like to thank my colleagues at ACTIAM N.V for their support and allowing me time to conclude this long journey.

Helmond, 29 August 2014

Michael van de Kerkhof

Table of Contents

1. Introduction

Sustainable and responsible investing, also known as social responsible investing (SRI) is a hot topic in the investment industry. From 2007 to 2010, SRI expanded at a higher pace than all other investment assets under professional management. SRI increased 13% whereas other investment assets increased merely 1% (Social Investment Forum Foundation, 2010).

SRI involves three key areas; ESG incorporation, shareholder advocacy and community investing. ESG is the most common part and involves screening firms on environmental, social and governance standards. There are two ways of screening; negative and positive. Negative screening excludes firms involved in "sin" activities such as tobacco, alcohol, nuclear power and weapons from an investors' universe. Positive screening ranks firms based on ESG criteria and often uses "best in class" approach to pick stocks that score highest on the investors' ESG criteria. As positive screening does not per se exclude firms that are involved in "sin" activities, investors often use a hybrid form of screening, not only excluding firms, but also ranking them in their efforts to be best in class.

As a result of growing interest for social responsible investing, it is useful the get more insight to the relationship of the screens that are used in relation to stock performances. Previous studies compared the performance of mutual funds and conventional funds in order to determine the profitablility of responsible investing. These studies do not account for managarial skills, which may determine the financial performance of these funds.

Furthermore, These studies take for granted that mutual funds differ in their selection criteria. Whereas most mutual funds are more orientated on environment and social performance, in general little attention is paid to the governance screen.

Other studies examine the performance of socially responsible investment portfolios by incorporating a single screen. While these studies are very useful in examining the relationship of that particular screen to investment performance, the interaction of other screens are not taken into account. In practice it seems that social responsible investors often use several screens at the same time. The Social Investment Forum (2010) indicates that the most important exclusionary screen that is used is the screen on the tobacco industry, whereas the environment screen is a very popular screen when implementing positive screening policies.

This thesis focuses on environmental, social and governance screens and their relation to stock performances. It tries to find linear relationships between the screens that are used and stock returns of the S&P500 over a time period of 1991 – 2012. Furthermore, this thesis takes controversial stocks into account when studying these relationships. Knowing the relationship between screens and stock performances enhances investment managers to make a profitable choice in screening criteria. Therefore the main question of the thesis is whether social responsible firms have better coefficients than non social responsible firms.

The main question is broken down into several hypotheses. First, it is research whether stocks that pass the negative screen threshold have a higher coefficient. Running a linear regression and performing a t-test shows that in fact the opposite is the case. Sin stocks seem to have a higher coefficient on stock returns than non controversial stock. Second, the environment, social and governance screens are studied by running linear regressions of these screens on stock returns for sample selections including and excluding sin stocks. The environment screen is not statistically significant it has a small negative coefficient. The coefficient for social screens is also negative, but it is statistically significant. Corporate governance screens are highly statistical significant and have a large positive coefficient. Third, regressions are performed on pairs of two screens at once, omitting one screen. In line

with previous findings, coefficients of environment and social screens rise when omitting the governance screen. When omitting the other variables, similar results appear when omitting environment and social screens. Fourth, the screens are tested for industry. No statistical significance was found, but in general, industries are positively related to the environment and negatively related to governance screens. Finally a time test is done for sub-periods 1991 – 2002 and 2003 – 2012. Including sin stocks shows a positive relation of environmental and social screens on return, but excluding sin stocks shows a positive relation of environmental and social screens on return.

This thesis separates itself from other literature as it grabs back to one of the most basic statistical research methods: linear regressions. Other studies use a more evolved model, such as CAPM or Carhart 4-factor model.

The rest of this paper is presented according to the following structure. Section two provides the current state of literature on social responsible investing and its relation to financial returns. In section 3 the hypotheses are formulated. The data that is used is described in section 4. Section 5 describes the methodology and the regressions that are used. The results are presented in section 6, while section 7 provides the conclusion. section 8 provides recommendations and limitations.

2. Current state of literature

2.1 Definitions

It is useful to clarify the definitions that are used in this thesis. As Schueth (2003) notes, there are many different terms that are used to describe the same phenomenon: *The terms social investing, socially responsible investing, ethical investing, socially aware investing, socially conscious investing, green investing, values-based investing, and mission-based or mission-related investing all refer to the same general process and are often used interchangeably.*" The terms that are interchangeably used in this thesis are socially responsible investing or responsible investing and are abbreviated as SRI.

In order to be on the same page, a few definitions are given of SRI. The forum for sustainable and responsible investment (2005) defines SRI as: "*an investment process that considers the social and environmental consequences of investments*". As it captures only the social and environmental angles, the definition of Renneboog, Ter Horst and Zhang (2008) may be more comprehensive: social responsible investing "*integrates social, environmental, and ethical considerations into investment decision making by applying a set of investment screens to select or exclude assets based on ecological, social, corporate governance or ethical criteria, and often engaging in the local communities and in shareholder activism to further corporate strategies towards above aim.*" This definition is a complete definition of SRI, however it is too broad within the framework of this thesis. The definition of Schueth (2003) may be more applicable as he defines SRI as: *"The process of integrating personal values and societal concerns into investment decision-making."* Furthermore he notes that there are three basic strategies in order to fulfill the dual goal of doing financially well and morally good. These three strategies are: screening, shareholder advocacy and community investing. As this thesis will focus only on the screening part, the definition of social responsible investing which is used in

this thesis can be defined as: *"The process of integrating personal values and societal concerns into investment decision-making, using different screening techniques."*

Screening is the most common form of social responsible investing and involves screening firms on environmental, social and governance (ESG) standards. In addition, screening can be split up into two different forms: negative and positive screening.
Negative screening involves the selection process of eliminating firms from the investment portfolio that do not satisfy the criteria of an investor. The firms that often are excluded are firms that are involved in morally controversial business activities such as tobacco, alcohol, nuclear power and weapons. Firms that are involved in these business activities are also known as "sin" stocks.
Positive screening is the process which ranks firms based on ESG criteria. It ranks firms that have a high score on an E,S or G criteria on top of the list. Furthermore it often distinct different industries as some industries are per definition for example more polluting. Investors often use a best-in-class approach to pick stocks that score highest on these ESG criteria, or to pick the top 40% stock on these ranks. As positive screening does not per se exclude firms that are involved in controversial activities, investors often use a hybrid form of screening. Hybrid screening does not only rank firms in their efforts to be best-in-class according to positive screening, but also exclude firms based on negative screening results. Negative screening leads to exclusion of firms of a social responsible investors' universe. The permitted firms within a portfolio for investors and fund managers is therefore smaller and constraints the diversification of the portfolio. In addition, positive screening also leads to exclusion of firms as a result of firm rankings. High ranked firms are included in investors' portfolios, whereas low ranked firms are typically excluded from the portfolio. The next paragraph will elaborate on this phenomenon.

2.2 Investment universe

Most papers discuss the positive and negative effects of screening on financial performance. Although both positive and negative effects are discussed, most papers emphasize the negative effects of social responsible investing or try to prove that it does not create value. Kurtz (2013) mentions that portfolio managers may decide to exclude sin stock, because it is cheaper to do so. His assumption is that a more complex research paradigm would lead to a more complex implementation. Having a list of firms that should be avoided would make the portfolio construction easier and cheaper. Creating a list of tobacco firms and excluding them from the portfolio is a very easy way of implementing negative screening with simple and widely accepted decision rules. Investments in firms that harm an individuals' health may be a more general and widespread definition. It is more complex to generate a list of those firms, because such a list is not generally accessible or easily created. In order to retrieve such a list, portfolio managers should have an agreement with an external research team or data providers, have an internal ESG research team, or both. An internal or external research team or a data provider clearly will add costs to the stock selection procedure. Depending on the level of ESG commitment, research will have to be done on the firms that comply with selection criteria or are currently on the edge of entering or falling out of the investment universe based on those criteria. Dialogue needs to be started with these firms in order to prevent them from falling back into bad habits or encourage them to increase their ESG commitment. As a result, a clear and simple selection rule may be cheap but as the list of exclusions becomes more complex, such a list becomes more expensive. These costs have to be made up for by investment funds, which often results into higher expense ratios.

Investor's may exclude firms based on their values and hence may not optimally diversify their portfolio. The large disinvestment of US institutional investors from firms

involved in South Africa was one of the first large stock boycotts in history. Many studies were conducted on the performance implications of excluding stocks. Andrew Rudd (1979) studied the impact of excluding South Africa investments and found that it leads to an annual underperformance of four basis points compared portfolio's which did not exclude South Africa. Weinstein, Alam and Blose (1991) performed an event study on US firms that had business relations to South Africa. They concluded only a very small or, after adjusting the event window, no price penalty was applied to these firms. Similar results were found by Teoh, Welch and Walsh (1997) who studied the increase in institutional holdings of firms after their divestments from South Africa. Neither the value of these firms, nor the financial markets in South Africa seem to be economically or statistically significantly affected by the boycott. In addition to Rudd (1979), Grossman and Sharpe (1986) show that different exclusion criteria have different impacts. Based on the criteria of Sullivan Principles, only a small set of firms are excluded which results in little effect on portfolio characteristics and performance. A very strict policy of excluding all firms that are related to South Africa leads to divestment in large cap firms such as Exxon Mobile, IBM, General Motors and Ford. Although excluding these large cap firms does negatively affect portfolio performance, investment in small cap firms leads to an outperformance of the NYSE by 0,187 percent. This suggests that social investors invest more in small cap firms. In order to investigate this phenomenon a study on the characteristics of assets held, portfolio diversification and variable effects of diversification on investment performance was performed by Bello (2005). He compared the characteristics of the socially responsible mutual funds with those of conventional funds and found that no significant differences are present. Furthermore, portfolio constraints neither seem to cause SRI funds to invest in fewer or smaller firms than do their conventional counterparts, nor does it allow for different investment performances.

An interesting study on the cost of SRI constraints on the return of a portfolio of US domestic mutual funds has been performed by Geczy, Stambaugh and Levin (2005). They made a case of an investor who wants to create an optimal portfolio in US domestic mutual funds with the highest Sharpe ratio based on ESG principles. The counterpart is an optimal portfolio of unscreened mutual funds. The difference between the returns of these portfolios shows the cost of SRI constraints. Three models are tested and investors who relied on the CAPM model only bear a cost of one or two basis points per month. Investors who relied on the Fama and French Three factor model incurred a cost of at least 30 basis points. When using the Carhart four factor model, costs are even slightly higher than 30 basis points a month. The most extreme case is where investors believe in managerial skills and rely on the track record of funds, believing that results in the past are a good estimator for future performance. They bear a cost of a stunning 1000 basis points. Clearly this is an extreme case, but the study does hint that implementing ESG criteria does add costs. Moreover, the authors state investors are too focused on the mean return and do not pay enough attention to the shape of return distribution. Their study clearly showed that the screened portfolio has a relative lack of positive outliers. For a skillful investor it would be harder to generate returns because of this smaller opportunity set.

Another way of studying the costs of SRI may be risk exposure of responsible investing. A company with high responsibility scores may bear lower litigation costs. A study on the risk exposure of the Domini Social Index and the S&P 500 index from 1990 to 1999 has been done by DiBartolomeo and Kurtz (1999). A crucial finding is that social screening would lead to a portfolio which differs from the S&P 500 on an economic and sector level. Hence the return of the social portfolio is implicitly related to its market exposure. The authors have back tested a risk-matched social portfolio on the S&P 500 and found that its return would be 1,49% per month versus 1,55% per month for the S&P 500. The authors find

the results to be similar, but the economic impact of six basis points a month, hints towards an even higher cost than the indicated one or two basis points by Geczy et. al. (2005).

A different approach on measuring the cost of SRI has been studied by Diltz (1995). He has researched the social screening impact on portfolio returns by constructing portfolio pairs. Dividing the pairs on basis of high versus low rated screen scores on ESG criteria of the Council on economic priorities, allows creating two different portfolios: one with high ranked scores and one with low ranked scores on the same ethical criteria. He concluded that screening has little impact on the return of a portfolio. Taking nuclear and military involvement into account, large differences are found between portfolios that are social responsible and those that are not. Portfolios that are favorably towards these unethical involvements have a higher alpha and are statistical significant on a 5% level in a two tail test. This raises questions on the ability of SRI to earn a premium over businesses involved in unethical industries, so called sin stocks.

2.3 Sin stocks

A study performed by Kacperczyk and Hong (2009) suggests that the South Africa boycott, is very specific, which may or may not be supported by other countries and regulations. Sin stocks are however are widely avoided by institutional investors. Kacperczyk and Hong (2009) studied the performance of sin stocks over a period from 1980 to 2006. The performance of a strategy in going long in sin stock excluding tobacco, and shorting comparables leads to an annual outperformance of 2.5%. Tobacco is excluded because these firms may increase litigation risk and may have unexpectedly positive earnings from litigation and therefore may drive the high returns of sin stocks. Angel and Rivoli (1997) agree that the expected returns on sin stocks should be higher than those of conventional stocks due to

limited risk sharing. Limited risk sharing increases idiosyncratic risk and therefore increases litigation risk.

Another argument for higher expected returns is the capital structure of sin stocks. Kacperczyk and Hong (2009) find that sin stock holdings by institutional investors are indeed lower than holdings in similar stocks. Chava (2011) supports these findings and adds that fewer banks are willing to provide loans to sin firms than to other firms. As investors shun sin firms the market to book value of these firms is lower than their comparables. In order to fund their operations, sin firms have a higher debt to equity ratio than comparables (Kacperczyk and Hong, 2009). In case of default equity investors are junior to debt investors, which results in higher required returns for equity investors. Kacperzyk and Hong (2009) suggested that firms that are active in the sin industry would therefore fund their activities with more debt than equity. Furthermore, institutional investors and mutual funds have to report their equity holdings on a semi-annual basis due to transparency rules. These transparency rules do not apply to corporate bond holdings, making it difficult to determine the debt investors. The result of Kacperzyk and Hong's research suggests that sin firms have a 19,3 percent higher leverage ratio than their comparables. The leverage effect results in a higher required expected return on equity.

Statman and Glushkov (2008) find that excluding sin stocks does hurt portfolio performance. Their study of stock returns in the period 1992-2007 illustrates that the shunning of stocks based on the criteria set by Kinder, Lydenberg and Domini (from here on quoted as KLD), leads to a return disadvantage compared to their conventional counterparts. This result is supported by Kempf and Osthoff (2007). They used the KLD database from 1992-2004 and studied the impact of several screens. They define a negative screen portfolio as a portfolio consisting of firms that are involved in at least one controversial business area and find that

negative screening erodes return. Statman and Glushkov (2008) however, also find that investing in stocks of firms that have high scores on social responsible criteria leads to an increase of portfolio return which largely offsets the exclusion of sin stocks. Hence positive screening seems to pay off.

2.4 Positive selection

A lot of research is done on the influence of positive screening on the portfolio return. Most studies use a high versus low comparison to determine whether financial performance of top performing firms outperform firms that are in the bottom of the rankings. In alignment with Statman and Glushkov (2008) Kempf and Osthoff (2007) found similar results. They created three portfolios based on negative screening, positive screening and a hybrid form using both negative and positive screening. The negative screening portfolio consists of firms that are involved in at least one controversial business area, whereas the positive screening portfolio includes the top ten percent of average positive ratings. Finally, the hybrid portfolio was created by first eliminating all controversial stocks based on the negative screening policy and then including the top ten percent of the remaining stocks. Using the Carhart four factor model, a strategy of going long in high SRI rated stocks and shorting the low SRI stock, would lead to a performance of five percent over the time period of 1992 to 2004.

Tsoutsoura (2004) studied the relation between corporate social performance (CSP) and corporate financial performance (CFP) for firms in the S&P 500 in the period of 1996 - 2000. Using KLD data and the DSI 400 index as a proxy for social firms, the result is that the relationship between CSP and CFP is positive and statistically significant. Moreover, KLD data provides stronger results than the DSI 400 data, confirming the hypothesis that KLD data is more sophisticated than DSI 400 index as a proxy. The study however does not check for causality.

Orlitzky, Schmidt and Rynes (2003) address to this causality issue in their meta-analysis paper. Although social performance seems to affect financial performance, the opposite also seems to be the case. Hence CSP may contribute to CFP, but financial successful firms also tend to spend more on social responsibility, causing a virtuous cycle.

Other studies have examined socially screened portfolio's based on one of the ESG factors. Konar and Cohen (2001) studied the market value of firms in the S&P 500 in relation to their environmental performance. They found that bad environmental performance is negatively correlated with the intangible asset value of firms. The authors conclude that legally omitted toxic has a large impact on intangible asset value. A 10% reduction in emissions of toxic chemicals results in a $34 million increase in market value. In line with these findings Derwall, Guenster, Bauer and Koedijk (2005) examined the impact of environmental screens by ranking firms according to the economic value they create relative to the waste they create. Two portfolios were constructed, one with high ranked and one with low ranked eco-efficient stock. The result is a significant outperformance of the high ranked portfolio, meaning that environmental screening would have a positive impact on portfolio performance.

Smith (1996) has examined the benefits of shareholder activism by performing a case study of CalPERS actions. CalPERS (California Public Employees' Retirement System) is regarded as a leader of activism in the US. He concludes that changing a firms' governance structure has statistically significant positive influence on shareholders wealth. Gompers, Ishii and Metrick (2003) find similar results when they constructed a governance index and using a strategy of investing in corporations with strong shareholder rights and going short on firms with weak shareholder rights. A follow up study has been done by Core, Guay and Rusticus (2006) as to whether there is a causal relationship between shareholder rights and a firms'

performance. They tested if operating cash flow differences caused by governance was unexpected by investors and found that surprises on earnings announcements do not explain differences in observed stock returns between firms with different shareholder rights. Therefore weak governance does not lead to weak stock performance.

Social screen studies were performed by Bello (2005), Hong and Kacperczyk (2009), Edmans (2011) and Derwall, Koedijk and Ter Horst (2010). Bello (2005) examines the differences in characteristics of assets held, portfolio diversification and effects of diversification on portfolio performance. He finds that "not a single characteristic of socially responsible mutual funds is significantly different from that of conventional funds." Edmans (2011) found that firms with high employee satisfaction earned a higher return than industry benchmarks.

2.5 Mutual funds studies

Hamilton, Jo and Statman (1993) suggest that social responsible investing may earn a premium as a result of mispricing social responsibility. This may be the case when the lower end tail of the return distribution is underestimated for negative news, which may result in class actions and therefore a higher litigation cost. Their research involved 32 social responsible mutual funds and 320 conventional funds. These numbers were split up into funds that were established before 1985 and after, as a lot of social responsible funds emerged in the 80's. Excess returns are calculated using Jensen's alpha. The average excess returns of both the 15 and 17 funds are not statistically different from zero. However, there is a lot of discrepancy between the best and the worst performing fund established in 1985 or earlier. Transamerica Capital Appreciation has an excess return of 5,74% per annum, whereas SFT Environmental awareness has an annual excess return of -6,33%, which leads to a difference of 12,66%. Furthermore, not all mutual funds have the same criteria to include a stock. They show that most funds include environmentally sound firms, but differ on other criteria, such

as employee criteria and social screens. This research does not look into the impact of the selection criteria of funds. The mutual fund of Dreyfus Third Century does include a lot of social screens and has an annual return of -4,01%, whereas Calvert Social Investment Fund seems to emphasize more on environmental screens and has an annual return of 0,08%. This may suggest that social screens seem to have a larger downward impact on return than environmental screens.

Bauer, Koedijk and Otten (2005) used a dataset in the period of 1990-2001 from the US, UK and Germany. They used both Jensen's alpha model and Carhart multifactor model to test whether returns on ethical investments transcends market cycles and style preferences. From the Jensen's alpha model the result is that the difference between the ethical and conventional alpha is statistically insignificant. A second result is that ethical funds tend to be less market sensitive than conventional funds. This result is supported by the test that was run with the Carhart multifactor model. Because this model consists of four model factors, it is expected to estimate mutual fund returns better than the CAPM model. This is confirmed as R^2 estimates are higher. An interesting result comes from the small minus big factor, which indicates that UK and German ethical mutual funds tend to be more oriented towards small cap, whereas US funds tend to be tilted more towards large caps. The fourth and final result comes from the high minus low market to book value. Ethical funds are more weighted towards growth stocks than conventional funds, which are more weighted towards value stocks. Overall findings are that mutual funds adjusted risk returns do not differ from conventional funds, even after correcting for management fees. Another test is performed on development of relative performance of ethical to conventional funds through time. The main result is that the performance of SRI mutual funds caught up on conventional funds. In the period of 1990-1993, ethical funds were trailing conventional funds. Mixed results are found from 1994-1997, where international oriented ethical funds seem to trail their conventional

counterparts, but are outperforming in the domestic market. From 1998 – 2001 all ethical fund markets are outperforming conventional funds.

A study performed by Kurtz and DiBartolomeo (1999) was done in order to determine if the Domini Social Index was really outperforming the S&P 500, or that its outperforming was a result of its sector and economic exposure. Using an extended CAPM model, also known as the Fundamental Risk model, and the Arbitrage Pricing model to minimize stock specific risk, the mean monthly return of the DSI outperformed S&P 500 by 0,18 percent over the period May 1990 to January 1999. Of this 0,18 percent outperformance, 0,06 percent was attributable to a higher beta of the DSI index. From the remaining 0,12 percent, 0,1 percent was coming from differences in industry exposure and only 0,02 percent arose from fundamental portfolio characteristics such as firm size and financial leverage. Results are significant at the 95 percent confidence level. The DSI has had a greater exposure to growth stocks, which implicitly lead to higher performance. For example, the Russell 1000 Growth index outperformed the Russell 1000 Value index by 0,23 percent. Abramson and Chung (2000) studied the possibility for social responsible investors to earn a premium by investing in value stock. Using a buy and hold strategy for the value stock portfolio, results in an outperformance of the benchmarks. Based on their own research from 1999, Kurtz and DiBartolomeo (2009) extended the range of data to June 2010. The industry factor is still positively related to the return and statistically significant using KLD400 index for the period of 1992 to 1999. From 2000 until 2010, the industries exposure was negatively related to the return and not statistically significant. Kempf and Osthoff (2007) found statistical significant outperformance by positive selection screens. However, when the observation period is divided into two sub periods (1992 – 1997 and 1998 -2004) the outperformance loses its statistical significance for both periods. Benson, Brailsford and Humphrey (2006) showed that SRI funds exhibit different industry betas consistent with different portfolio positions, but that

these differences vary from year to year. Estimated industry betas between the two groups are significantly different for the telecommunications and utilities industries, which indicate that portfolios of SRI and conventional investors do in fact differ. It is also found that there is little difference in stock-picking ability between the two groups of fund managers.

Renneboog, Ter Horst and Zhang (2008) compared SRI mutual funds in the US, UK, continental Europe and Asia-Pacific countries to their domestic benchmarks. They found that the SRI mutual funds underperformed their benchmarks for the majority of countries. In France, Japan and Sweden no statistical difference was found between SRI and conventional mutual funds. SRI investors show some ability in avoiding stocks that will perform poorly, but cannot pick winners. Furthermore, the screening policy of SRI funds has a significant impact on financial performance. Corporate governance and social screens have a negative influence on risk-adjusted returns, whereas community investing and employing an in-house SRI research team have a positive effect on the financial performance. Barnett and Salomon (2006) found similar results when comparing 61 mutual funds over a time period of 1972 to 2000. Their research illustrates that financial performance suffers when implementing social screens at first, but as screening intensifies, financial performance increases. Environmental and social screens seem to be costing the most performance, whereas community investing improves performance.

3. Hypotheses

The study by Hong and Kacperczyk (2009) illustrated that sin stocks are underpriced and outperform comparables. Hence, neglecting these stocks has an influence on the cost of capital of large institutional investors. Derwall et al. (2010) have studied the returns of shunned stocks and positive social and environmental screens. Controversy stock which are shunned, have a higher return than comparable stock. As positive screening also provides higher returns, the effect of excluding controversy stock and positive screening would cancel each other out. This raises the question whether the performance contribution of the sin stock is partly integrated into the return of the positive screens? As stated earlier, positive screening does not necessarily exclude sin stocks. Perhaps a part of the positive screening return is a result of investments in sin stocks. In order to investigate this phenomenon I first study the impact of sin stocks on portfolio returns. The first hypothesis is:

H1: Stocks that pass the negative screen threshold have a higher coefficient

If sin stocks in fact provide us with higher returns than SRI stocks, it may be that the responsible investing comes with a cost. Because many papers have shown different outcomes for different screens, it useful to study the relationship of different screens on stock performances. Testing for environmental, social and corporate governance screens individually may pinpoint which screening criteria are best implemented or best avoided when creating portfolios. Derwall, et al. (2005), Konar and Cohen (2001) find that environmental screens improve portfolio performance, whereas Hamilton and Statman (1993) do not find any evidence of outperformance of stocks with high environment scores. Social screens seem to have a both positive and negative influence on portfolio returns. Bello (2005) and Benson, et al. (2006) find that there is no significant difference in performance of highly

ranked social stock and conventional stocks. Then again, Orlitzky, et. al (2003) and Statman (2000) find that social screening does add value. Finally, the governance screen seems to be studied less than the environment and social screens. Nonetheless, Core, et al. (2006) did not find any causal relationship between shareholder rights and a financial performance. A result which was found by Gompers et al. (2003) was that stocks with strong shareholder rights outperformed stocks with low shareholder rights. A literature overview on several studies that research different ESG criteria is added in the appendix A.

As research shows contrasting results, it is useful to test every screen independently for both sin and non sin stocks. Therefore the hypotheses two to four, in nil form, are:

H2: Stocks that pass the negative screen threshold have a higher environment coefficent..
H3: Stocks that pass the negative screen threshold have a higher social coefficent..
H4: Stocks that pass the negative screen threshold have a higher governance coefficent..

Kurtz and DiBartolomeo (2009), showed that the it is useful to see investment performances of responsible investing over a long time period. They showed that the KLD400 index outperformed the S&P500 from its inception in 1992. However, since 2000 the outperformances are reversed and it has been the S&P500 which outperformed the KLD400 index until 2008. The KLD400 index consists of 250 S&P500 firms and 150 firms that are selected on their outstanding responsible scores or because they represent under-presented industries. According to Kurtz and DiBartolomeo (2009) the index was overweighted in the technology sector, which has driven the outperformance in the 1990's. As the index remained overweighted in the technology sector, it comes as no surprise that dot com bubble hurt performance early in the 2000's. It is argued that some industries are more subject to negative screening than others. A firm which has its main activities in the oil industry is likely to receive a lower environment score than a financial would receive (Statman and Glushkov,

2008). Therefore, it may be useful to study the impact of an industry on the ESG score of a firm. In order to study the impact of the industry regressions on the coefficients of several screens, hypothesis five is:

Another part of research is not about the screens itself, but on the test sample.

H5: The industry a stock is in influences ESG coefficients.

Next to the industry test, a time test is performed. As shown by Kurtz and DiBartolomeo (2009), the performance of the DS400 has drop in the 2000's, where it first outperformed the S&P 500. Kempf and Osthoff (2007) find no performance differences when dividing their data period into two sub-periods. In order test the data for time effects, hypothesis 6 is:

H6: Coefficients are higher for the before 2003, than there after.

4. Data and methodology

4.1 Databases

The data is retrieved from Wharton Research Data Services. First all historical constituents in the 1991 to 2012 period of the S&P500 index are obtained, using Compustat. Furthermore Compustat was used to obtain the historical prices of the S&P500 index .

Second, all historical prices of the S&P500 constituents for the period of 1991 to 2012 are retrieved using Center for Research in Security Prices (CRSP) monthly S&P500 Index Constituents. Also the Standard Industrial Classification codes are taken from this database. Each year the index may alter due to corporate actions or as a result of S&P500 criteria such as market cap and sector representation. Therefore, the constituents which have not been a part of the S&P500 for at least six months in a single year are eliminated. If a company would be incorporated into the S&P500 for only 2 months, the data from this company would be dubious as a yearly return.

Third, the screening data is obtained from the Kinder, Lydenberg and Domini (KLD) database. The period of 1991 – 2012 is chosen, because I want to utilize all available screening data that is available. The KLD database is a binary database. Firms are evaluated on many criteria. These criteria are divided into strengths and concerns. For each strength or concern either a one or a zero is graded. In case of a strength (e.g. support for housing, charitable giving), a rating of one indicates that the firm is believed to have fulfilled the KLD requirements for this particular screen. A zero is graded in absence of criteria fulfillment. The concern criteria are assigned in a similar way. A score of one indicates that a firm has that particular concern (e.g. hazardous waste, substantial emissions) and a zero indicates that the concern is not relevant for this firm. Following the way firm is graded, firms with high strength scores and many zero concern scores will receive a high responsible investing score.

Likewise, Firms with many zero scores on strength indicators and high scores on concerns will receive a low responsible investing score. In this research the strength and concern scores are netted, meaning that scores of concerns are subtracted from the strength scores. It is important to realize that no weights are assigned to these screening criteria. Thus a score of one on charitable giving is equally weighted as a score on of one on recycling.

4.2 KLD data

For this research I retrieved as much screening data as possible for these stocks, meaning that I will include as many ESG criteria as possible. However, not all ESG criteria have been consistently evaluated throughout the time period. For example, the screen for South Africa is a very specific screen for firms that are involved in business in or with South Africa from 1991 – 1994. Although this screen is very specific and it covers a short period, I have decided to use every bit of data possible from the database. As a result, this data will be included in the 1991-1994 time periods, but will not be included in other years for the simple reason that there is no such data available. Hence the number of screens will vary each year. Table 1 illustrates the total number of screens for each year.

Table 1
Overview of the number of environmental, social and governance screens for 1991 - 2012

	Environment				Corporate Covernance		Social			Sin	
Year	Total	Negative screens	Positive screens	Total	Negative screens	Positive screens	Total	Negative screens	Positive screens	Sin screens	Total number of screens
1991	12	6	6	5	2	3	37	16	21	11	65
1992	12	6	6	6	3	3	27	12	15	11	56
1993	12	6	6	6	3	3	38	17	21	11	67
1994	12	6	6	6	3	3	46	21	25	11	75
1995	12	6	6	6	3	3	44	19	25	11	73
1996	11	6	5	6	2	4	32	16	16	11	60
1997	11	6	5	7	3	4	32	16	16	13	63
1998	11	6	5	7	3	4	44	20	24	15	77
1999	12	7	5	7	3	4	44	20	24	15	78
2000	12	7	5	7	3	4	46	21	25	15	80
2001	12	7	5	7	3	4	46	21	25	15	80
2002	12	7	5	7	3	4	46	20	26	6	71
2003	12	7	5	7	3	4	47	20	27	6	72
2004	12	7	5	7	3	4	47	20	27	6	72
2005	12	7	5	11	6	5	48	20	28	6	77
2006	13	7	6	11	6	5	48	20	28	6	78
2007	13	7	6	13	7	6	48	20	28	6	80
2008	13	7	6	12	6	6	48	20	28	8	81
2009	13	7	6	12	6	6	48	20	28	8	81
2010	13	7	6	6	4	2	38	16	22	6	63
2011	13	7	6	6	4	2	38	15	23	6	63
2012	18	9	9	8	5	3	38	18	20	6	70

The table provides us an overview of the large number of screens available and shows that not all factors have the same number of screens available to them. For instance, environment commences with 12 screens available, but changes during its time span to end up at 18 screens in 2012. These totals are a result of the change of the number of both negative and positive screens. Both positive and negative screens start of at 6 screens in 1991 and end up at 9 screens in 2012. However, as the number of negative screens only increases from 1991 on, the number of positive screens decreases from its original six screens to five screens from 1996 – 2005, before returning to six screen from 2006 – 2011 and ending at 9 screens in 2012.

Furthermore, the screens are not distributed evenly among positive and negative screens. In 1991 there are more positive than negative corporate governance screens, but in

2005 it is the other way around. Environment screens tend to be more negative than positive over the entire period, whereas social screens tend to be positive. Moreover as the awareness of social responsible investing seems to intensify, so does the number of screens. It seems that the number of screens increases until 2010, when some screens were merged or eliminated. The same pattern can be found for the sin screens. However, these screens were already cut back in 2002.

It is noticed that the number of social screens is substantially higher than for any other screen. The reason for this is that the social screen is a cumulative of several sub screens. Table 2 illustrates the screens and number of factors which are included into the social screen. An overview of all screens of the KLD database can be found in Appendix B.

Table 2

Overview of the subset of social screens and the number of social screens

	Community				Diversity			Employee Relations				Human Rights			Product	
Year	Total	Negative	Positive	Total	Negative	Positive	Total	Negative	Positive	Total	Negative	Positive	Total	Negative	Positive
1991	8	4	4	9	2	7	10	4	6	2	2	0	8	4	4
1992	8	4	4	9	2	7	0	0	0	2	2	0	8	4	4
1993	7	3	4	10	3	7	11	5	6	2	2	0	8	4	4
1994	10	4	6	10	3	7	11	5	6	7	5	2	8	4	4
1995	10	4	6	11	3	8	10	5	5	5	3	2	8	4	4
1996	10	4	6	0	0	0	10	5	5	4	3	1	8	4	4
1997	10	4	6	0	0	0	10	5	5	4	3	1	8	4	4
1998	10	4	6	11	3	8	10	5	5	5	4	1	8	4	4
1999	10	4	6	11	3	8	10	5	5	5	4	1	8	4	4
2000	10	4	6	11	3	8	10	5	5	7	5	2	8	4	4
2001	10	4	6	11	3	8	10	5	5	7	5	2	8	4	4
2002	10	4	6	11	3	8	10	5	5	7	4	3	8	4	4
2003	10	4	6	11	3	8	11	5	6	7	4	3	8	4	4
2004	10	4	6	11	3	8	11	5	6	7	4	3	8	4	4
2005	11	4	7	11	3	8	11	5	6	7	4	3	8	4	4
2006	11	4	7	11	3	8	11	5	6	7	4	3	8	4	4
2007	11	4	7	11	3	8	11	5	6	7	4	3	8	4	4
2008	11	4	7	11	3	8	11	5	6	7	4	3	8	4	4
2009	11	4	7	11	3	8	11	5	6	7	4	3	8	4	4
2010	4	1	3	10	3	7	10	4	6	6	4	2	8	4	4
2011	5	1	4	10	3	7	10	4	6	5	3	2	8	4	4
2012	3	1	2	6	3	3	14	5	9	6	4	2	9	5	4

The goal of this study is to find out whether firms with strong SRI ratings are rewarded for their policy and outperform firms with weak ratings. The KLD database is used to grade their policies. This database does not allow for any grey area on the scale of good or

bad. It is very resolute in grading a firm either a concern or strength. The disadvantage of this binary grading is that some firms that or on the edge of receiving a rating are harshly categorized and for instance, receive a zero where a score of one would be okay as well. Such black and white ratings may, on the one hand lead to "box-checking" by firms. Although it is doubtful for a firm to assign a minority to the board, who has not got any influence on the firms' policy, it is a possibility for firms to comply with the demands of KLD to receive a positive rating on the diversity screen. On the other hand, the KLD database gives a clear boundary and makes the data easy to work with. Furthermore, the database is widely used in numerous studies. In line with its strict interpretations, firms that have at least a score on one of the negative screens: alcohol, fire arms, gambling, military involvement, nuclear involvement or tobacco, are labeled 'sin' stocks. On the qualitative scale of the environmental, governance and social screens, firms may receive both positive and negative ratings. As mentioned in the database paragraph, the numbers of strength and concern screens vary, both in number and through time. For the environment screen it would not be unusual to find a mean that is negative, because there are more negative than positive screens. Exxon Mobile for example has one strength and four concern screens when it comes to the environment. Netting these numbers would mean that Exxon Mobile would score minus 3 points on the environment screen. But Exxon does not only have negative ratings, when it comes to human rights, it has two strength ratings and no concerns, meaning it would receive a score of plus two. Lastly, it may occur that the concern and strengths cancel each other out. Exxon Mobile got rated two concerns and two strengths in the employee relations screen in 2000. As these ratings cancel each other out, the score for Exxon Mobile would be zero. An overview of the summary statistics of the screens is given by table 3.

Table 3

Summary statistics of the overall netted total scores, strengths minus concerns, of environment, social and governance screens. It shows the number of observations, mean, standard deviation, median, minimum, maximum, median, first and last quartile for environment, social and governance screens for the period 1991-2012.

Variable	Obs	Mean	Std. Dev.	Min	Q25	Median2	Q75	Max
Environment	10471	-0.0895807	1.167613	-5	0	0	0	5
Social	10471	0.7597173	2.452557	-7	-1	0	2	14
Governance	10471	-0.4489543	0.7330684	-4	-1	0	0	2

Table 3 shows that the mean of both environment and corporate governance screen are negative, which indicates that firms in general do not meet SRI qualifications. This is not a surprise for the environment screen, as table 1 already showed us that there are more negative environment screens than positive environment screens. A very striking observation is that for all three screens, the median is 0. For the environment screen even the first quartile and fourth quartile are zero as well. This indicates that either a firm has no positive or negative scores, or that the number of positive and negative screens cancel each other out. When viewing the minimum and maximum scores of the companies individually, the minimum score that is given to a single firm is minus five, whereas the maximum score is five. This suggests that the environmental screen has some very high and very low outliers, but in general these would cancel each other out. For the governance screen, similar results can be found. However, the minimum and maximum score are less dispersed from minimum four to maximum two. Then again, the social screen has a wide range from minimum 7 to maximum 14. As the number of screens is much higher than for the other screens, it is no surprise that first and last quartile are not equal to zero. However, with a median of zero, again most companies tend to stick in the middle, doing neither harm nor good. Table 4 shows a more detailed distribution of scores for different percentiles.

Table 4

Summary statistics of the overall netted total scores for environment, social and governance screens. It shows the percentiles 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95% and 99%, minimum and maximum scores for environment, social and governance screens for the period 1991-2012.

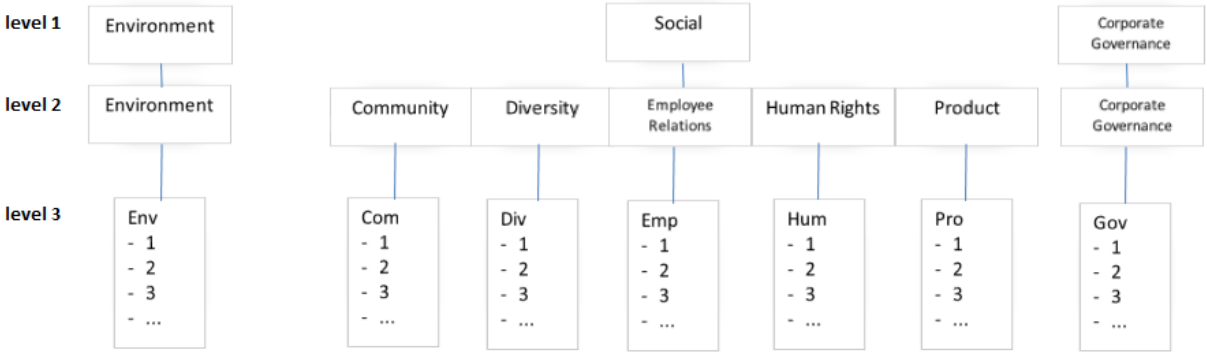
Screen	Min	P1	P5	P10	P25	P50	P75	P90	P95	P99	Max
Environment	-5	-4	-2	-1	0	0	0	1	2	3	5
Social	-7	-4	-3	-2	-1	0	2	4	5	8	14
Governance	-4	-2	-2	-1	-1	0	0	0	1	1	2

5. Methodology

5.1 Screen formation

Figure 1 illustrates the model which is used in this research and the different levels and criteria that are used to estimate the screens. Level 1 consists of three broad criteria: environment, social and corporate governance. It is the aggregated score of the level two screens. For environmental and governance, this is the same as the level 2 screen. The social screen is the aggregate of five sub screens: community, diversity, employee relations, human rights and product screens. The product screen is placed into the social screening, because it incorporates screens on antitrust, benefits to economically disadvantaged and customer relations. Level two is built of the aggregated scores of level three factors, which consist of multiple sub factors both positive and negative. For example, the five corporate governance level 3 factors in 1991 are: high compensation and ownership concern for the concern screens, while for the strength screens these are limited compensation, ownership strength and other strengths.

Figure 1: schematic overview of screen levels and factors.



A similar model is used for the negative screens. Level 1 is the labeled sin stocks and level 2 are the actual negative screens on one of the six negative screens: alcohol, fire arms, gambling, military involvement, nuclear involvement or tobacco. which are aggregated at level one. These screens include alcohol, fire arms, gambling, military involvement, nuclear involvement and tobacco. Each of these screens can be broken down into several sub screens. Nevertheless, as sin stocks are defined as stocks of firms that are involved in at least one controversial business activity, all level 3 factors with a concern will rate the stock a "sin stock".

5.2 Model estimations

This thesis studies the impact of screening policies on the investment performance of S&P500 stocks. The dependant variables will be the historical returns of the S&P500 stocks. The ESG variables which are used, will be used as independent variables. For each year the stocks are categorized as either sin stock or non sin stock. Running linear regressions allows examining whether a relation exists between social screens and stock returns. This thesis deviates from existing literature on the relationship of social responsible investing and financial performance. It does not value screening criteria nor does it differentiate S&P500 stocks as all stocks are treated equally weighted. This study attempts to find a linear relation of returns and screening criteria without using developed models such as the capital asset

pricing model, Fama-French three factor model or the Carhart four factor model. Furthermore it studies the relationships of all ESG criteria on financial performance rather than just one aspect.

Hypothesis one states that sin stocks have lower return than non controversial stocks. To test this hypothesis, a linear regression is performed over the entire period from 1991 - 2012. In model 1a (model 1b), the return of sin (non-sin) stocks is the dependant variable and a dummy variable indicating a sin (non-sin) stock is the independent variable.

Model 1a: $r_{sin} = \beta_0 + \beta_1 \text{ sin} + \varepsilon$

Model 1b: $r_{nonsin} = \beta_0 + \beta_1 \text{ nonsin} + \varepsilon$

Where r_{sin} (r_{nonsin}) is the return of (non-)sin stocks, β_0 is a constant, β_1 is the coefficient of (non-)sin stocks and ε is the residual.

In order to get a more refined touch of the performance of social responsible firms, two models are estimated. First I will estimate a model on all screens at level one; environmental, social and governance. This model includes firms that are marked as sin firms and is labeled model 2a. Then I will estimate model 2b, which is exactly the same as model 2a, but excludes sin stocks. If sin stocks have lower return than non controversial stocks, I would expect the betas of model 2a to be lower than model 2b betas.

Model 2a: $r_{sin} = \beta_0 + \beta_1 \text{ env} + \beta_2 \text{ soc} + \beta_3 \text{ gov} + \varepsilon$

Model 2b: $r_{nonsin} = \beta_0 + \beta_1 \text{ env} + \beta_2 \text{ soc} + \beta_3 \text{ gov} + \varepsilon$

Where r_{sin} (r_{nonsin}) is the return of (non-)sin stocks, β_0 is a constant, $\beta_1$env is the coefficient of the environment screen, β_2 soc is the coefficient of the social screen, β_3 gov is the coefficient of the governance screen and ε is the residual.

Zooming in on each screen independently, hypotheses two, three and four, study the impact of environmental, social and governance screens on the performance of the sample for the entire time period from 1991-2012. The regression that is run is based on the level one aggregated scores of each screen. Once again the samples differ, running linear regressions on the entire database for the models that are numbered as "a" and excluding sin stocks from the models that are numbered "b".

Model 3a: $r_{sin} = \beta_0 + \beta_1 \text{ env} + \varepsilon$

Model 3b: $r_{nonsin} = \beta_0 + \beta_1 \text{ env} + \varepsilon$

Model 4a: $r_{sin} = \beta_0 + \beta_1 \text{ soc} + \varepsilon$

Model 4b: $r_{nonsin} = \beta_0 + \beta_1 \text{ soc} + \varepsilon$

Model 5a: $r_{sin} = \beta_0 + \beta_1 \text{ gov} + \varepsilon$

Model 5b: $r_{nonsin} = \beta_0 + \beta_1 \text{ gov} + \varepsilon$

Where r_{sin} is the return of the portfolio including sin stocks, r_{nonsin} is the return of the portfolio excluding sin stocks, β_0 is a constant, β_1 env is the coefficient of environment screens, β_1 soc is the coefficient of social screens and β_1 gov is the coefficient of governance screens.

A model with two screens combined can be compared with a model which uses one or three screens and reveal more on the omitted variable. To get more conclusive results, six more models are estimated where two screens are combined. Similar to previous models, the "a" numbered models include sin stocks and its counter model excluding sin stock is numbered "b".

Model 6a: $r_{sin} = \beta_0 + \beta_1 \text{ env} + \beta_2 \text{ soc} + \varepsilon$

Model 6b: $r_{nonsin} = \beta_0 + \beta_1 \text{ env} + \beta_2 \text{ soc} + \varepsilon$

Model 7a: $r_{sin} = \beta_0 + \beta_1 \text{ env} + \beta_2 \text{ gov} + \varepsilon$

Model 7b: $r_{nonsin} = \beta_0 + \beta_1 \text{ env} + \beta_2 \text{ gov} + \varepsilon$

Model 8a: $r_{sin} = \beta_0 + \beta_1 \text{ soc} + \beta_2 \text{ gov} + \varepsilon$

Model 8b: $r_{nonsin} = \beta_0 + \beta_1 soc + \beta_2 gov + \varepsilon$

Where r_{sin} is the return of the portfolio including sin stocks, r_{nonsin} is the return of the portfolio excluding sin stocks, β_0 is a constant, β_1 env is the coefficient of environment screens, β_1 soc and β_2 soc is the coefficient of social screens and β_2 gov is the coefficient of governance screens.

Next, regressions on level two are performed. Although both environment and governance screens are similar to their screens at level one, the difference with previous models is that the social screen is split up into five different sub screens: community, diversity, employee relations, human rights and product. Drilling down on screens allows us to better understand which screens are more likely to add value and which ones are more likely to reduce value. Models 9a and 9b are linear regression models on the sub screens at level two and include environment and corporate governance screens. Models 10a and 10b exclude environment and governance screens and focuses on social screens only. Model 9a and model 10a include sin stocks, whereas model 9b and 10b exclude sin stocks.

Model 9a: $r_{sin} = \beta_0 + \beta_1 env + \beta_2 com + \beta_3 div + \beta_4 emp + \beta_5 hum + \beta_6 pro + \beta_7 gov + \varepsilon$

Model 9b: $r_{nonsin} = \beta_0 + \beta_1 env + \beta_2 com + \beta_3 div + \beta_4 emp + \beta_5 hum + \beta_6 pro + \beta_7 gov + \varepsilon$

Model 10a: $r_{sin} = \beta_0 + \beta_1 com + \beta_2 div + \beta_3 emp + \beta_4 hum + \beta_5 pro + \varepsilon$

Model 10b: $r_{nonsin} = \beta_0 + \beta_1 com + \beta_2 div + \beta_3 emp + \beta_4 hum + \beta_5 pro + \varepsilon$

Where r_{sin} is the return of the portfolio including sin stocks, r_{nonsin} is the return of the portfolio excluding sin stocks, β_0 is a constant and for model 9a and 9b β_1 to β_7 are the coefficient of environment, community, diversity, employee relations, human rights, product and governance screens respectively. For model 10a and 10b β_1 to β_5 are the coefficient of environment, community, diversity, employee relations, human rights and product.

5.3 Industry influences

As discussed by Statman and Glushkov (2008) Kurtz and DiBartolomeo (1999) and Benson, Brailsford and Humphrey (2006), the industry in which a firm is operating may influence the score of a particular screen. As argued, oil companies are expected to have a lower score on environment screens than financials. In order to study the impact of industry on the coefficients of the screens, all firms in the dataset are coded using the Standard Industrial Classification. The firms are categorized into 10 categories, which were taken from http://siccode.com. Table 5 shows the industries and the number of firms in these industries.

Table 5

Summary statistics of the industries in which firms operate for the period 1991-2012. SIC group is the industry group a firm belongs to. SIC number is the number which is given to the industry. Frequency shows the number of observations of firms in this industry. Percentage shows the percentage of the industry in the entire database.

SIC group	SIC number	Frequency	Percentage
Agriculture, Forestry, Fishing	1	2	0,02
Mining	2	527	5,03
Construction	3	93	0,89
Manufacturing	4	4.878	46,59
Transportation & Public Utilities	5	1.317	12,58
Wholesale & Trade	6	236	2,25
Retail trade	7	803	7,67
Finance, Insurance, Real Estate	8	1.658	15,83
Services	9	956	9,13
Public administration	10	1	0,01
Total		10.471	100

From table 5 it is clear to see that the manufacturing industry is by far the largest industry and that public administration and agriculture are the smallest. Because these two industries have so little contribution, they are omitted from the models that estimate the influence of the industry on a particular screen.

Model 11a: $env = \beta_0 + \beta_1 2 + \beta_2 3 + \beta_3 4 + \beta_4 5 + \beta_5 6 + \beta_6 7 + \beta_7 8 + \beta_8 9$

Model 11b: $soc = \beta_0 + \beta_1 2 + \beta_2 3 + \beta_3 4 + \beta_4 5 + \beta_5 6 + \beta_6 7 + \beta_7 8 + \beta_8 9$

Model 11c: $gov = \beta_0 + \beta_1 2 + \beta_2 3 + \beta_3 4 + \beta_4 5 + \beta_5 6 + \beta_6 7 + \beta_7 8 + \beta_8 9$

Where env is the environment screen, soc is the social screen, gov is the governance screen, β_0 is a constant and β_1 to β_8 are the coefficients of the industries as labeled in table 5.

5.4 Time period

As discussed by Kurtz and DiBartolomeo (2009), it is interesting to see investment performances of responsible investing over a long time period. They studied the long term performance of KLD400 social investment index over a time period of 1992-2010. An interesting result is that the KLD400 index outperformed the S&P500 from its inception in 1992. However, since 2000 the outperformances are reversed and it has been the S&P500 which outperformed the KLD400 index until 2008. Although based on almost the same database, Kempf and Osthoff (2007) used KLD ratings which includes S&P500 and another 150 DS400 stocks, no differences between the sub-periods 1992 – 1997 and 1998 – 2004 were found. The methodology used is a long short strategy, going long in top 10% of socially rated stocks and going short in bottom 10% of socially rated stocks.

In this study the time period of 1991 – 2012 is divided into two time spans: 1991 – 2002 and 2003 – 2012. This time span is chosen as it seperates the entire time period into two more or less equal tome spans. Therefore models 12a and 12b are estimated as:

Model 12a: $r_{sin} = \beta_0 + \beta_1 env + \beta_2 soc + \beta_3 gov + \varepsilon$

Model 12b: $r_{nonsin} = \beta_0 + \beta_1 env + \beta_2 soc + \beta_3 gov + \varepsilon$

Where r_{sin} is the return of the portfolio including sin stocks, r_{nonsin} is the return of the portfolio excluding sin stocks, β_0 is a constant, $\beta_1$env is the coefficient of the environment screen, β_2 soc

is the coefficient of the social screen, β_3 gov is the coefficient of the governance screen and ε is the residual. These models are ran for time periods 1991 – 2002 and 2003 – 2012.

6. Results:

6.1 Sin stocks and ESG overview

The output in table 5a illustrates the results of the regressions using model 1a and 1b. Results are statistically significant and induce an outperformance of sin stocks on non sin stocks of 42 basis points. As expected, β_1 sin and β_1 non-sin are each other's opposites, because the data is set as a dummy 1 when stocks are sin and zero when stocks are non-sin.

Table 6a

Output of regression models 1a and 1b.
Model 1a regresses return on sin stock and Model1b regresses return on non sin stock. The returns of sin and nonsin stocks are given in percentages. The test was performed using robust standard errors to account for heteroskedasticity. *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

Variable	Model 1a	Model 1b
Sin	0.210***	
	(0.0660)	
Nonsin		-0.210***
		(0.0660)
Constant	1.027***	1.237***
	(0.0331)	(0.0571)
N	10471	10471
R^2	0.001	0.001
adj. R^2	0.001	0.001
F	10.13	10.13

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To see whether the mean returns of sin stocks really are higher than the returns of non sin stocks, a t-test is ran. Note that sin stocks are indicated using a "1" and non sin stocks are

noted "0". Table 6b shows the results. Again results show that sin stocks have a higher mean than non controversial stocks. Even more, the absolute value of the t-statistic is larger than 2, which indicates it is significant at a 95% confidence level.

Table 6b
Output of two sample t-test with equal variances. Sin stocks are in group 1 and non controversial stocks are in group 0.
The table shows the number of observations, mean return, standard error, standard deviation and 95% confidence interval of sin (1) and nonsin (0) stocks. It also shows the combined and the difference of mean returns and standard error., as well as 95% confidence interval. H_0 shows the tested hypothesis, also 3 alternative hypotheses are given: difference < 0, Difference not equal to 0 and difference > 0. The degree of freedom is shown and the t-statistic is given by t.

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	8409	.0102666	.0003311	.0303598	.0096176	.0109156
1	2062	.0123672	.0005711	.0259334	.0112472	.0134872
combined	10471	.0106802	.0002888	.0295513	.0101142	.0112463
diff		-.0021006	.0007259		-.0035236	-.0006776

diff = mean(0) - mean(1) t = -2.8937
Ho: diff = 0 degrees of freedom = 10469

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 0.0019 Pr(|T| > |t|) = 0.0038 Pr(T > t) = 0.9981

Next the more elaborated models 2 to 5 are performed. Table 7 shows the results of the tested models. Model 2 includes environment, social and governance screens and only the governance screen is highly statistical significant at both model 2a and 2b. The difference is 0,1 basis points, which is neglectable. However, the economic impact of the governance screen seems to be substantial with 17 basis points and contributes the most to the stock performances. The environmental screen is not statistically significant, but hints that responsible firms sacrifice more performance in return for good environmental scores than their not responsible counterparts. The social screen is statistically significant at the 1% level for sin stocks, but not for the responsible stocks in both models 2 and 4. The coefficient

indicates that social screens are negatively related to returns. Moreover, sin stocks are more vulnerable to social screening than responsible stocks.

In line with model 1, environmental screens show that responsible firms underperform non responsible firms, but that the underperformance is not statistical significant. The social screen has a positive effect on performance, but is not statistical significant in model 2a and 2b. In model 4a and 4b the social screens has a negative contribution to performance. The result is significant at the ten percent level.

Table 7

Output of regression models 2 to 5, which regress the return on the different screening criteria. Models 2a and 2b, 3a and 3b, 4a and 4b, and 5a and 5b are paired and similar to each other. The models which are named a include sin stocks and the models named b exclude sin stock. Estimates are noted in percentages. The tests were performed using robust standard errors to account for heteroskedasticity *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

Variable	Model 2a	Model 2b	Model 3a	Model 3b	Model 4a	Model 4b	Model 5a	Model 5b
Environment	-0.0169	-0.0297	-0.0315	-0.0363				
	(0.0229)	(0.0288)	(0.0219)	(0.0272)				
Social	-0.0372***	-0.0224			-0.0355***	-0.0230		
	(0.0135)	(0.0156)			(0.0129)	(0.0148)		
Governance	0.173***	0.172***					0.161***	0.163***
	(0.0435)	(0.0502)					(0.0434)	(0.0501)
Constant	1.172***	1.127***	1.065***	1.026***	1.095***	1.046***	1.140***	1.104***
	(0.0331)	(0.0388)	(0.0290)	(0.0331)	(0.0306)	(0.0359)	(0.0308)	(0.0357)
N	10471	8409	10471	8409	10471	8409	10471	8409
R^2	0.003	0.002	0.000	0.000	0.001	0.000	0.002	0.002
adj. R^2	0.002	0.002	0.000	0.000	0.001	0.000	0.001	0.001
F	7.982	5.055	2.073	1.784	7.552	2.399	13.77	10.59

6.2 ESG factors

In addition to models that are based on either one screen independently or all screens together, it might be useful to look at the combination of two screens in a single model. Excluding one variable from the regression allows us to compare the performance of the remaining categories to the variable that is left out. The results of these combinations can be found in table 8. Leaving the governance screen out of model 6 shows no real influence on the environmental and social variable. Both variables are still negatively related to performance, but not as strong as in the previous models. However, compared to the 0,173 coefficient in model 2a, the return of eliminating the governance screen in model 6a would only earn about 1 basis point, which means a reduction of 16 basis points. This result also applies to the responsible models 2b and 6b.

Excluding social screens from model 2 leads to model 7. When comparing the results of these two models, the negative impact of the environment variable on return seems to be worse by almost 1 basis point. Additionally, the environment variable has become statistically significant at the 10 percent level, whereas the governance variable is still statistically significant at the one percent level. Although its contribution to performance has dropped about a half basis point, the combined effect on the environment and governance variable is lower than the effect of the social screen as given in model 2. Model 2a showed a coefficient of -0,0372 where the combined impact of removing the social screen in model 7a is roughly -0,095. This would mean that eliminating the social screen would lead to an additional 2,8 basis points. Surprisingly, for model 2b and model 7b this would only lead to a positive contribution of one and a quarter basis point.

Finally in model 8, the environment variable is dropped. Although the variable had little economic significance and no statistical significance in neither of the models, dropping it

would only cause marginal effects for the social and governance screen. Therefore the contribution of dropping the environment variable would increase return by 1,5 basis points.

Dropping variables is easy in theoretical models, but is more difficult in practice. Eliminating a screen would mean that a firm would have a netted score of zero on this variable. For some firms this might actually be the case, but if not than earning a strength point is likely to bear costs. nevertheless, the models indicate what firms should aim for; high corporate governance scores. The large impact of the governance screens may stern from the fact that only few screens are available as seen in table 1. As a result, increasing the governance score by one point is not such an easy task.

Table 8

Output of regression models 6, 7 and 8, which regress the return on the different screening criteria. Models 6a and 6b, 7a and 7b, and 8a and 8b are paired and similar to each other. The models named a include sin stocks and the models named b exclude sin stock. The tests were performed using robust standard errors to account for heteroskedasticity *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

Variable	Model 6a	Model 6b	Model 7a	Model 7b	Model 8a	Model 8b
Environment	-0.0101	-0.0221	-0.0399*	-0.0456*		
	(0.0228)	(0.0287)	(0.0220)	(0.0273)		
Social	-0.0340**	-0.0198			-0.0396***	-0.0266*
	(0.0135)	(0.0157)			(0.0129)	(0.0148)
Governance			0.166***	0.168***	0.171***	0.169***
			(0.0435)	(0.0502)	(0.0434)	(0.0501)
Constant	1.093***	1.043***	1.139***	1.106***	1.175***	1.130***
	(0.0311)	(0.0362)	(0.0308)	(0.0357)	(0.0328)	(0.0386)
N	10471	8409	10471	8409	10471	8409
R^2	0.001	0.000	0.002	0.002	0.003	0.002
adj. R^2	0.001	0.000	0.002	0.002	0.002	0.002
F	3.967	1.593	8.395	6.593	11.62	6.991

Next, a regression on level 2 is performed in order to get more insight on the contribution of the 5 social sub screens. First models 9a and 9b are estimated. These models include environment, social and governance screens and respectively include and exclude sin stocks. Then models 10a and 10b are estimated, including only the social screens on level2: community, diversity, employee relations, human rights and product. Table 9 shows the results.

Table 9

Output of regression models 9a, 9b, 10a and 10b, which regress the return for the 1991 – 2012 period on the different screening criteria on level 2; environment, community, diversity, employee relations, human rights, product and governance. Models 9a and 9b include environment and corporate governance screens. Model 9a includes sin stocks and model 9b excludes sin stocks. Model 10a and 10b are performed on exclusively the social screens. The tests were performed using robust standard errors to account for heteroskedasticity *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

Variable	Model 9a	Model 9b	Model 10a	Model 10b
Environment	-0.0182	-0.0251		
	(0.0233)	(0.0295)		
Community	0.128***	0.124***	0.133***	0.127***
	(0.0350)	(0.0417)	(0.0343)	(0.0404)
Diversity	-0.140***	-0.125***	-0.147***	-0.134***
	(0.0224)	(0.0253)	(0.0222)	(0.0250)
Employee Rel.	-0.0160	0.0194	-0.0118	0.0231
	(0.0308)	(0.0352)	(0.0308)	(0.0353)
Human rights	-0.195***	-0.222**	-0.194***	-0.227***
	(0.0708)	(0.0862)	(0.0693)	(0.0846)
Product	0.0462	0.0706*	0.0601*	0.0846**
	(0.0345)	(0.0406)	(0.0339)	(0.0399)
Governance	0.132***	0.124**		
	(0.0438)	(0.0507)		
Constant	1.192***	1.145***	1.143***	1.097***
	(0.0371)	(0.0424)	(0.0359)	(0.0407)
N	10471	8409	10471	8409
R^2	0.007	0.006	0.006	0.005
adj. R^2	0.007	0.006	0.006	0.005
F	9.785	6.916	12.61	8.957

From this table it can is shown that the community, diversity, human rights governance and to some extent product screen are all statistically significant. Examining the

models tells that diversity and human rights have a large negative impact on a the performance, ranging from -12,5 basis points to -22,7 basis points. For the human rights variable, this is no real surprise. Table 2 showed that there are more negative than positive screens for human rights. In 1991 and 1992 no positive screens were available while screens on South Africa and Northern Ireland were negative screens from 1991 until 1994. The diversity variable does not have a particular screen that clarifies the negative coefficient. On the contrary, there are more positive than negative diversity over the entire sample period. Therefore it was expected that the diversity variable would contribute to performance rather than reduce performance. As expected from previous models, the environment variable is not significant and the governance variable is positively related to return and statistically significant. The community variable is also statistically significant at the 1% level and adds more than 12 basis points to a firms' return. This positive relationship is expected, because more positive screens are available than there are negative screens.

6.3 Industry influences

Several conclusions can be drawn from table 9. First, No industry shows statistical significance for any of the screens. Second, no single industry is positively related to all three screens. The Mining industry is the only industry which has a negative coefficent on all three screens. Third, the environment screen is positively related to all industries, except for the mining industry. It has an average positive coefficient of about 0.6 which is economically significant. Fourth, all industries are negatively related to the social screen. Coefficients variate from as low as -2.129 in the construction industry to as high as -0.0275 in the financial industry. Fifth, the coefficients for the governance screens are all negative.

Table 10

Output of regression models 11a, 11b and 11c, which shows the beta coefficients of industries on screening policies. The tests were performed using robust standard errors to account for heteroskedasticity *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

	Environment	Social	Governance
Mining	-0.0810	-2.052	-0.332
	(0.660)	(1.394)	(0.420)
Construction	0.667	-2.129	-0.634
	(0.669)	(1.412)	(0.425)
Manufacturing	0.550	-0.440	-0.402
	(0.659)	(1.391)	(0.419)
Transportation & Public Utilities	0.213	-1.005	-0.325
	(0.659)	(1.392)	(0.419)
Wholesale & Trade	0.679	-0.956	-0.449
	(0.663)	(1.399)	(0.421)
Retail Trade	0.880	-0.802	-0.410
	(0.660)	(1.393)	(0.419)
Finance, Insurance & Real Estate	0.829	-0.0275	-0.593
	(0.659)	(1.391)	(0.419)
Services	0.856	-0.359	-0.688
	(0.660)	(1.392)	(0.419)
Constant	-0.667	1.333	4.63e-11
	(0.659)	(1.390)	(0.418)
N	10471	10471	10471
R^2	0.047	0.037	0.024
adj. R^2	0.046	0.036	0.023
F	63.79	50.16	31.55

6.4 Time period regressions

Following Kurtz and DiBartolomeo (2009) and Kempf and Osthoff (2007) a time study is performed over two sub periods which are divided by the dot com bubble in 2002. Table 10 shows the results of linear regression models 12a and 12b.

Table 11

Output of regression models 12a and 12b, which shows the beta coefficients of different screening policies and time spans. The left panel shows the results for the 1991 – 2002 time span and the right panel displays the results for the 2003 – 2012 time span. Model 12a is ran on all firms and model 12b is ran only on non exclusionary stocks. The tests were performed using robust standard errors to account for heteroskedasticity *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

	Time period 1991 - 2002		Time period 2003 - 2012	
	Model 12a	Model 12b	Model 12a	Model 12b
Environment	0.102**	-0.0717**	0.134**	-0.0992**
	(0.0406)	(0.0356)	(0.0524)	(0.0431)
Social	0.000723	-0.0560***	0.0339	-0.0497***
	(0.0185)	(0.0169)	(0.0223)	(0.0191)
Governance	0.316***	0.0991*	0.354***	0.0964
	(0.0640)	(0.0532)	(0.0740)	(0.0605)
Constant	1.268***	1.113***	1.208***	1.099***
	(0.0468)	(0.0578)	(0.0553)	(0.0664)
N	5735	4736	4452	3957
R^2	0.006	0.005	0.008	0.005
adj. R^2	0.005	0.004	0.007	0.004
F	10.65	7.429	11.29	6.253

Time period 1991 – 2002 shows some striking differences when it comes to including or excluding sin stocks. For model 12a all screening policies are positively related to return, with exception for the social screen which is nor statistical nor economically significant. When excluding sin stocks, model 12b shows that only the governance screen is stil positively related to return. However, it has lost most of its economic impact, contributing only 10 basis points, whereas model 12a showed a coefficient of 31 basispoints. Next to the reduced governance factor, the coefficients of both environmental and social screens have changed from having a positive into having a negative relation to returns. The time period of 2003 – 2012 shows similar results, however these results seem to have a stronger effect. The environment screen has not only a higher coefficient in model 12a for the 2003 – 2012 time span, but also a lower coefficient in model 12b.

As a result of these regressions, in order to study the impact of sin stocks on the coefficients the same regressions are ran but only for sin stocks. The result is not what was

expected. Because the hybrid model 12a which includes all stocks has higher coefficients than the clean model 12b which exempt sin stocks, it was expected that a regression on sin stock only would show positive coefficients. Instead only the social screen is statistically significant and has a negative coefficient. Therefore it may be useful to further investigate the relationship of sin stocks on non exclusionary stocks.

Table 12

Output of regression model 12a and model 12b, which shows the beta coefficients of different screening policies and time spans for sin stocks. The left panel shows the results for the 1991 – 2002 time span and the right panel displays the results for the 2003 – 2012 time span. The tests were performed using robust standard errors to account for heteroskedasticity *N* is the number of observations R^2 and adj.R^2 show the model fit and F shows the degree of freedom. The standard errors are in parentheses. *, **, *** indicate the significance level at 10%, 5% and 1% respectively.

	Time period 1991 -2002	Time period 2003 – 2012
	Sin stocks	Sin stocks
Environment	0.0941	0.00270
	(0.0621)	(0.0578)
Social	-0.0822**	-0.0749**
	(0.0319)	(0.0356)
Governance	0.141	0.113
	(0.126)	(0.109)
Constant	1.437***	1.190***
	(0.0863)	(0.111)
N	1283	779
R^2	0.008	0.006
adj. R^2	0.005	0.002
F	3.228	1.608

7. Conclusion:

Sustainable and responsible investing, also known as social responsible investing (SRI) expanded at a higher pace than all other investment assets from 2007 to 2010. This raises the question on how SRI is best implemented. This study analyzes different screening policies based on KLD database. The sample is restricted to the S&P 500 stocks for the period of 1991 – 2012. Instead of using models such as CAPM or the Carhart 4-factor model, linear regressions are ran in order to determine whether such relations exists between screening policies and returns. The main results are inconclusive. Sin stocks seem to have a higher coefficient on stock returns than non controversial stock. When regressing all three screens on return, the environment screen is not statistically significant it has a small negative coefficient. The coefficient for social screens is also negative, but it is statistically significant. Corporate governance screens are highly statistical significant and have a large positive coefficient. When performing pairs of two screens at once, omitting one screen, coefficients of environment and social screens rise when omitting the governance screen. Similar results appear when omitting environment and social screens. The industry test shows that no statistical significance was found, but in general, industries are positively related to the environment and negatively related to governance screens. Finally a time test is done for sub-periods 1991 – 2002 and 2003 – 2012. Including sin stocks shows a positive relation of environmental and social screens on return, but excluding sin stocks shows a positive relation of environmental and social screens on return.

8. Recommendations and limitations:

This thesis uses linear regression models to study the relationship of screens and stock returns of the S&P 500. A basic regression only can do so much. It is recommended that other modeling methods are tested, especially since the results are inconclusive. Due to time restrictions, this thesis was done by using data that was easily available. The KLD database is widely used by researchers, but other databases may provide a good alternative to study screening relations for the US or other markets.

References

Angel, J.J. and Rivoli, P. (1997). Does ethical investing impose a cost upon the firm? A theoretical -perspective. *Journal of Investing,* Vol. 6, no. 4, pp. 57-61.

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

Bello, Z.Y. (2005). Socially responsible investing and portfolio diversification. *Journal of financial research,* Vol. 28, no. 1, pp. 41-57.

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,* Vol. 61, no. 2, pp. 655-687.

Derwall, J., Guenster, N., Bauer, B., and Koedijk, K. (2005). The eco-efficiency premium puzzle. *Financial analysts journal,* Vol. 61, No. 2, pp. 51-63.

Diltz, J.D., (1995). Does social screening affect portfolio performance? *Journal of Investing,* Vol.4, No. 1, pp. 64-69.

Derwall, J., Koedijk, K., and Ter Horst, J. (2010). A tale of values-driven and profit-seeking social investors. *Journal of banking and finance,* Vol. 35, No. 8, pp. 2137-2147.

Gompers, P., Ishii, J., and Metrick, A. (2003). Corporate Governance and equity prices. *The Quarterly Journal of Economics*, Vol. 118 no. 1, pp.107-155.

Grossman, B. R. and W. F. Sharpe (1986). Financial implications of South African divestment, *Financial*
Analysts Journal, vol. 42, pp. 15–29.

Hamilton, S., Jo, H., and Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds, *Financial analyst journal,* Vol. 49, No. 6, pp. 62-66.

Hong, H., and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics,* Vol. 93, pp. 15-36.

Kempf, A., and Osthoff, P. (2007). The effect of Socially Responsible Investing on Portfolio Performance. *Journal of European Financial Management,* Vol. 13. No. 5, pp. 908-922.

Konar, S., and Cohen, M.A. (2001). Does the market value environmental performance? *The review of economics and statistics,* Vol. 83, No. 2, pp. 281-289.

Renneboog, L., Ter Horst, J., and Zhang, Z. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance,* Vol. 32, pp. 1723-1742.

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

Philips, Hager and North Investment Management (2007) *Does SRI Hurt Investment Returns?* Available on https://www.phn.com/portals/0/pdfs/Articles/20071012DoesSRIHurtInvestmentReturns.pdf

Rudd, A. (1979). Divestment of South African equities: How risky? *Journal of Portfolio Management*, Vol. 5, pp. 5-10.

Statman, M. and Glushkov, D. (2008). The wages of responsibility.

Social investment forum foundation. (2005). *Report on socially responsible investing trends in the United States*. Available on http://ussif.org/resources/pubs/index.cfm

Social investment forum foundation. (2010). *Report on socially responsible investing trends in the United States*. Available on http://ussif.org/resources/pubs/index.cfm

United Nations Environment Programme Finance Initiative & Mercer. (2007). *Demystifying responsible investment performance*. Available on http://www.unepfi.org/publications/investment/index.html

Appendix A

Table I: overview of literature studies. Authors and publication year in brackets are shown in the first column, the title of the study is in the second column, the time period of the study is in the third column, the sort of screen that is researched is in column 4; where E is invironment, S is social and G is governance. Column 5 shows the if the findings are P, positive, NS, neutral or N, negative. The last column briefly summerizes the research.

	Authors	**Title of study**	**Time period of study**	**E, S or G**	**Findings on ESG factors**	**Results**
1	Abramson, L. & Chung, D. (2000)	Socially responsible investing: Viable for value investors?	Sep 1990 - Mar 2000	ESG	P	Based on DSI a growth and a value stock portfolio are built. The value portfolio is simulated using a buy and hold strategy whereas the growth portfolio is rebalanced quarterly. The result is that both portfolios outperform the benchmarks over the studied time period.
2	Barnett, M. & Salomon, R. (2006)	Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance.	Jan 1972 - Dec 2000	E and S	NS - P	Financial performance suffers when implementing social screens at first, but as screening intensifies, financial performance increases. Environmental and social screens seem to be costing the most performance, whereas community investing improves performance.
3	Bello, Z. (2005)	Socially responsible investing and portfolio diversification.	Jan 1994 - Mar 2001	S	NS	A study on the characteristics of assets held, portfolio diversification and variable effects of diversification on investment performance. Comparing the characteristics of the SRI mutual funds with those of conventional funds showed no significant differences. Furthermore, portfolio constraints neither seem to cause SRI funds to invest in fewer or smaller companies

						than do their conventional counterparts.
4	Benson, K.L., Brailsford, T.J. & Humphrey, J.E. (2006)	Do socially responsible fund managers really invest differently?	Jan 1994 - Dec 2003	S	NS	This paper examines whether the portfolio allocation across industry sectors and the stock-picking ability of SRI managers are different when compared to conventional fund managers. The study finds that SRI funds exhibit different industry betas consistent with different portfolio positions, but that these differences vary from year to year. It is also found that there is little difference in stock-picking ability between the two groups of fund managers.
5	Core, J., Guay, W. & Rusticus, T. (2006)	Does weak governance cause weak stock returns? An examination of firm operating performance and investors' expectations.	Sep 1990 - Dec1999	G	NS	A study on the causal relationship between shareholder rights and a firms' performance. They tested if operating cash flow differences caused by governance was unexpected by investors and found that surprises on earnings announcements do not explain differences in observed stock returns between firms with different shareholder rights. Weak governance does not lead to weak stock performances.
6	Derwall, J., Guenster, N., Bauer, R. & Koedijk, K. (2005)	The eco-efficiency premium puzzle.	Jul 1995 - Dec 2003	E	P	This study examines the performance impact of using environmental ratings as part of an active management strategy. Creating a high and a low rated environmental portfolio and

						comparing differences resulted in benefits of considering environmental criteria in the investment process can be substantial.
7	Geczy, C., Stambaugh, R. & Levin, D. (2005).	Investing in Socially Responsible Mutual Funds	Jul 1963 - Dec 2001	S	N	The study researches the costs of investing in SRI mutual funds compared to conventional funds. The overall result is that when investors believe in the CAPM pricing model, and not in manager skill, then the costs of SRI constraints is small. However, as investors believe in multi-factor models or manager skill, the costs of SRI constraints increase.
8	Hong, H. & Kacperczyk, M. (2009)	The price of sin: The effects of social norms on markets	Jan 1965 - Dec 2004	S	N	The authors find that 'sin' stocks are underpriced and outperform comparables, and that this might be explained by social norms that restrict ownership by more 'visible' institutional investors and insurance companies.
9	Orlitzky, M., Schmidt, F.L. & Rynes, S.L. (2003)	Corporate social and financial performance: A meta-analysis	Jan 1972 - Dec 1997	S	P	The authors find that there is a positive relationship between Corporate Social Performance (CSP) and Corporate Financial Performance (CFP), but the extent of the positive correlation is impacted by characteristics such as reputation, disclosure of CSP, and market measure of CFP.
10	Statman, M. (2000)	Socially responsible mutual funds.	May 1990 - Sep 1998	S	P	Comparing Domini Social Index (DSI) with S&P 500 performance showed a higher risk-adjusted return of DSI compared to S&P 500. When comparing 31 SRI mutual funds to S&P 500, SRI mutual funds underperform, but

						outperformed similar conventional mutual funds.
11	Konar, S. & Cohen, M.A. (2001)	Does the Market Value Environmental Performance?	1989	E	P	A study the market value of S&P 500 firms in relation to their environmental performance. Bad environmental performance is negatively correlated with the intangible asset value of firms. The authors conclude that legally omitted toxic is negatively related to and has a large impact on intangible asset value.
12	Gompers, P., Ishii, J. & Metrick, A. (2003)	Corporate Governance and equity prices	1990, 1993, 1995 & 1998	G	P	Investing in firms with strong shareholder rights and shorting stocks of firms with low ratings would have earned abnormal return of 8,5%
13	Renneboog, L., Ter Horst, J. & Zhang, C. (2008)	The price of ethics and stakeholder governance: The performance of socially responsible mutual funds	Jan 1991 - Dec 2003	ESG	NS - N	Compared mutual funds worldwide to their domestic benchmarks and found that most SRI funds underperform their domestic benchmark. Social and governance screens yield lower risk-adjusted returns.
14	Hamilton, S., Jo, H. & Statman, M. (1993)	Doing well while doing good? The investment performance of socially responsible mutual funds	Jan 1981 - Dec 1990	ES	NS	A study of SRI mutual funds versus conventional funds, which concludes that there are no differences in risk-adjusted returns between SRI and conventional funds.
15	Bauer, R., Koedijk, K. & Otten, R. (2005)	International evidence on ethical mutual fund performance and investment style	1990 - 2001	ESG	NS	This study finds no evidence of significant differences in risk-adjusted returns between ethical and conventional funds in the US, UK and Germany.

16	Kurtz., L & DiBartolomeo, D. (1999)	Managing Risk Exposures of Socially screened portfolios	May 1990 - Jan 1999	ESG	NS	Show that the return differences between the Domini Social Index and S&P 500 do not arise from the socially responsible behavior of the included companies, but from economic and sector exposures that are the implicit result of social screening of portfolio securities.
17	Kempf, A. & Osthoff, P. (2007)	The Effect of Socially Responsible Investing on Portfolio Performance	1991 - 2003	ESG	P	A research to study the impact of social screens on financial performance. A strategy of buying stocks with high socially responsible ratings and selling stocks with low socially responsible ratings leads to high abnormal returns of up to 8.7% per year

Appendix B

Overview of the KLD screens that are used in this thesis.

Negative screens	Concern
Tobacco	Tobacco Involvement
Tobacco	Tobacco Other Concern (through 2002)
Military involvment	Military Involvement
Military involvment	Minor Weapons Contracting (1991-2002)
Military involvment	Major Weapons-related Supplier (1991-2002)
Military involvment	Military Other Concern (through 2002)
Nuclear Involvement	Nuclear Involvement
Nuclear Involvement	Nuclear Design (through 2002)
Nuclear Involvement	Nuclear Fuel Cycle (through 2002)
Nuclear Involvement	Nuclear Other Concern (through 2002)
Firearms	Firearms Involvement (from 1999)
Gambling	Gambling Involvement
Gambling	Gambling Other Concern (through 2002)
Alcohol	Alcohol Involvement
Alcohol	Alcohol Other Concern (through 2002)
Alcohol	Alcohol - Number of Concerns

Positive screens	Strength	Concern
Environment	Beneficial Products and Services	Hazardous Waste
Environment	Pollution Prevention	Regulatory Problems
Environment	Recycling	Ozone Depleting Chemicals
Environment	Clean Energy	Substantial Emissions
Environment	Property, Plant, Equipment (through 1995)	Agriculture Chemicals
Environment	Management Systems Strength	Climate Change (from 1999)
Environment	Environment Other Strength	Negative Impact of Products and Services
Environment	Water Stress	Supply Chain Management
Environment	Biodiversity & Land Use	Water Management
Environment	Raw Material Sourcing	Land Use & Biodiversity
Environment		Non Carbon Releases
Environment		Environment Other Concerns
Corporate governance	Limited Compensation	High Compensation
Corporate governance	Ownership Strength	Ownsership Concern

Corporate governance	Transparency Strength	Accounting Concern (from 2005)
Corporate governance	Political Accountability Strength (from 2005)	Transparency Concern (from 2005)
Corporate governance	Public Policy Strength (from 2007 through 2011)	Political Accountability Concern (from 2005)
Corporate governance	Corp. Gov Other Strength	Public Policy Concern (from 2007 through 2011)
Corporate governance	Controversial Investments	Governance Structures Controversies
Corporate governance	Business Ethics	Corp. Gov Other Concerns
Corporate governance	Corruption & Political Instability	
Corporate governance	Financial System Instability	
Community	Charitable Giving (from 1991 through 2011)	Investment Controversies
Community	Innovative Giving	Negative Economic Impact
Community	Support for Housing	Tax Disputes
Community	Support for Education (from 1994)	Community Other Concerns
Community	Non-US Charitable Giving	
Community	Volunteer Programs (from 2005)	
Community	Community Engagement	
Community	Other Strengths (from 1991 through 2011)	
Diversity	CEO	Controversies
Diversity	Promotion (from 1991 through 2011)	Non-Representation (from 1993 through 2011)
Diversity	Board of Directors	Board Diversity
Diversity	Work-Life Benefits (from 1991 through 2011)	Diversity Other Concerns
Diversity	Women and Minority Contracting	Board of Directors - Minorities
Diversity	Employment of the Disabled	
Diversity	Gay and Lesbian Policies (from 1995 through 2011)	
Diversity	Employment of Underrepresented Groups	
Diversity	Diversity Other Strength	
Employee Relations	Union Relations	Union Relations
Employee Relations	No-Layoff Policy (through 1994)	Health and Safety Concern
Employee Relations	Cash Profit Sharing	Workforce Reductions
Employee Relations	Employee Involvement	Retirement Benefits Concern
Employee Relations	Retirement Benefits Strength	Supply Chain Controversies
Employee Relations	Health and Safety Strength	Emp. Relations Other Concerns
Employee Relations	Supply Chain Policies, Programs & Initiatives	Child Labor

Employee Relations	Employee Strengths - Other Strengths (from 1991 through 2011)	
Employee Relations	Compensation & Benefits	
Employee Relations	Employee Relations	
Employee Relations	Professional Development	
Employee Relations	Human Capital Management	
Human Rights	Positive Record in S. Africa (1994-1995)	South Africa (1991-1994)
Human Rights	Indigenous Peoples Relations Strength (from 2000)	Northern Ireland (1991-1994)
Human Rights	Labor Rights Strength (from 2002)	Burma Concern (from 1995)
Human Rights	Human Rights Other Strength	Mexico (1995-2002)
Human Rights		Labor Rights Concern (from 1998)
Human Rights		Indigenous Peoples Relations Concern (from 2000)
Human Rights		Operations in Sudan (from 2010 through 2011)
Human Rights		Human Rights Other Concerns
Human Rights		Freedom of Expression & Censorship
Human Rights		Human Rights Violations
Product	Quality	Product Safety
Product	R+D-Innovation	Marketing-Contracting Concern
Product	Benefits to Economically Disadvantaged	Antitrust
Product	Access to Capital	Product Other Concerns
Product	Product Other Strengths	Customer Relations