UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

BACHELOR THESIS IN FINANCIAL ECONOMICS
AT THE DEPARTMENT OF ECONOMICS, SPRING 2014

Superior Fund Performance by Exclusion

Does an Exclusion and Norms-Based Strategy Enhance Performance for Socially Responsible Retail Investors?

Authors:
Mikaela Hildén
Fredrik Olsson Stenkil

Supervisor:
Charles Nadeau

ABSTRACT

The study examines whether there is a significant relationship between risk-adjusted returns and the intensity of exclusion and norms-based screening. The approaches to socially responsible investing adopted by mutual funds in the Swedish premium pension system are quantified and the monthly performances over the three year period March 2011 – March 2014 are estimated with Jensen's alpha, Carhart's four-factor alpha and the Sharpe ratio. The results show significant advantages of applying nine out of twelve applicable screens to the socially responsible investment strategy when Jensen's alpha and the Sharpe ratio are measurements of performance. The relationship is quadratic with a diminishing positive effect and is ascribed to a trade-off between Portfolio Theory and Stakeholder Theory.

We wish to thank our supervisor Charles Nadeau, Senior Lecturer at Gothenburg School of Business, Economics and Law, Department of Economics, for his valuable guidance and insight used in the process of writing this thesis.

Keywords: Socially Responsible Investing, Mutual Fund Performance, Environmental, Social and Governance (ESG) Analysis, Ethical Investments, Exclusion- and Norms-Based Screening, Sustainability

CONTENTS

1. INTRODUCTION

1.1. Background

Since its first broad introduction in the 1960's, the public awareness of sustainability has steadily increased as documented by The International Institute for Sustainable Development (2012). The trend is largely spread over a variety of business areas and many leading organisations and leaders of the world focus extensively on environmental issues and possible sustainable solutions. A common verdict is that something has to be done to reduce the depletion of non-renewable resources and somehow the costs of externalities have to be compensated. The message is expanding to the financial world and the Generation Foundation, by many considered being the frontrunner sustainable investment fund proclaimed "sustainability research integrated into a rigorous traditional investment process will deliver superior long term results". Chairman Al Gore further stated, "Integrating issues such as climate change into investment analysis is simply common sense". However, despite these fine words there is scepticism towards socially responsible investing (SRI) and the question remains whether these investments can deliver superior risk-adjusted returns.

SRI is a topic difficult to quantify and research has shown that there are substantial challenges in performing SRI studies. A common belief has developed that socially responsible investing often leads to inferior returns and compared to conventional investing the investment style has therefore not been looked at as an attractive investment choice, as stated by Shröder (2004) and Bauer, Koedijk & Otten (2005). However, researchers from Maastricht University, Riedl & Smeets (2014), argue that the reasons for choosing socially responsible investing could be explained by factors other than only financial performance and they define responsible investing as individually biased in terms of expected return and risk perception. In fact, Bollen (2007) found that investors are likely not to hold on to badly performing conventional mutual funds while socially responsible investors tend to hold on to badly performing SRI mutual funds. Previous research indicates that people do care about other elements except for financial performance, which makes it clear that SRI is associated with irrational economic utility maximisation and varying individual preferences.

In the study "Social Identification and Investment Decisions" by Bauer & Smeets (2013), evidence was found that portfolio choice is driven by intrinsic social values and that social reputation might be a factor affecting investment choices. The asset selection process is linked to individual investor preferences more than solely to risk-return expectations; altruism and social norms are thus proven to be powerful elements in decision-making (Fehr & Fischbacher, 2003). There are retail investors who put value in avoiding investing in unethical or irresponsible companies. SRI can thus be desirable without expecting superior returns if the investor considers the reduced negative externalities to be worth the price. However, investors without such strong preferences for SRI might also pursue responsible investments if the financial benefits are strong enough.

Stuart Hart, Professor of Management at Cornell University, argues that doing business globally is clearly in a changing phase and he anticipates that economic and industrial change will be driven by sustainable development over the next 25 years. If the prediction is realised, it is important to develop the research and investigate methods for evaluating SRI mutual fund performance. Essentially, SRI is "ethical investments, responsible investments, sustainable investments, and any other investment process that combines investors' financial objectives with their concerns about environmental, social and governance (ESG) issues" as defined by Eurosif, the European Sustainable Investment Forum. ESG Managers®, a Morningstar initiative, has put together a summary of ESG companies' activities and potential economic gains, shown in Table 1.

There are typically two types of goals in ethical investing, the economic rational goal of wealth-maximisation and the goal of taking social responsibility. According to Eurosif's European SRI Study from 2012, the six most commonly used approaches for SRI are exclusion of companies based on certain criteria e.g. specific sectors, excluding based on breaches against international norms, engaging and influencing regarding ESG matters, selecting only best-in-class investments based on ESG criteria, integrating ESG factors into the financial analysis and finally, investing with the purpose of developing sustainability. The Swedish market has been on the forefront of socially responsible investing for many years and data on the Swedish market breakdown by SRI strategy, shown in Table 2, illustrates that the focus lies on exclusion and norms-based screening.

Table 1: ESG companies' activities and gains, ESG Managers® (2014)

Area of focus	Activity	Potential impact on financial performance
Environment	- Resource management and pollution prevention - Reduced emissions and climate impact - Environmental reporting/disclosure	- Avoid or minimise environmental liabilities - Lower costs/increase profitability through energy and other efficiencies - Reduce regulatory, litigation and reputational risk - Indicator of well-governed company
Social	**Workplace** **-** Diversity **-** Health and safety **-** Labour-Management relations **-** Human rights **Product Integrity** - Safety - Product quality - Emerging technology issues **Community Impact** - Community relations - Responsible lending - Corporate philanthropy	**Workplace** - Improved productivity and morale - Reduce turnover and absenteeism - Openness to new ideas and innovation - Reduce potential for litigation and reputational risk **Product Integrity** - Create brand loyalty **-** Increase sales based on products safety and excellence **-** Reduce potential for litigation **-** Reduce reputational risk **Community Impact** - Improve brand loyalty - Protect license to operate
Corporate Governance	- Executive compensation - Board accountability - Shareholder rights - Reporting and disclosure	- Align interests of shareowners and management - Avoid negative financial surprises or "blow-ups" - Reduce reputational risk

Table 2: Swedish market breakdown by SRI strategy, Eurosif (2012)

SRI Strategy	Millions of SEK	Percentages
Exclusion Screening	3 030 607	39.59 %
Norms-Based Screening	2 313 365	30.22 %
Engagement and Voting	1 227 928	16.04 %
Best-In-Class	768 314	10.04 %
Integration of ESG	311 284	4.07 %
Sustainability Themed	3 532	0.05 %

1.2. Purpose and Contribution

The purpose is to study the historical relationship between mutual funds' intensity in their SRI approach and their risk-adjusted return. Drawing conclusions as to what the optimal level of exclusion and norms-based screening intensity was during the period March 2011 – March 2014 could help establish a reference base for further research. Future research may become extensive enough to be used as indicators for predictions, providing useful guidelines for retail investors wishing to understand the impact on returns from differences in SRI approaches between mutual funds.

The study 'Beyond Dichotomy...' by Barnett & Salomon (2006) emphasises that SRI mutual funds are not homogenous, and if a foundation is to be laid out for investors aiming at investing responsibly, Capelle-Blancard & Monjon (2012) states *"The debate on the economic and financial impact of corporate social responsibility should move away from the question "does it pay to be good?" to "when does it pay to be good?"*

The contribution of this study is to highlight and quantify the varying characteristics of SRI practices on the Swedish market in order study the impact of common SRI approaches, providing guidance as to the optimal screening intensity for maximising performance.

1.3. Research Questions

The question of interest is whether the number of exclusion and norms-based screening criteria employed by fund managers had an impact on risk-adjusted performance during the sample period. The reason behind focusing on negative screening is because exclusion and norms-based screening criteria make up almost 70 % of the market divided by strategy as previously shown in table 2, and is thus the most widely applied method for Swedish SRI fund managers.

Drawing inspiration from previous studies, allowing for a better fitting model and realising that there might be a turning point in the SRI screening process where its effect either starts to be realised or begins to decline, a curvilinear relationship is the main hypothesis, resulting in the following three hypotheses.

I. *There is a quadratic relationship between the number of exclusion and norms-based screening criteria and risk-adjusted return, when measured as the estimated Jensen's alpha.*

II. *There is a quadratic relationship between the number of exclusion and norms-based screening criteria and risk-adjusted return, when measured as the estimated Carhart's four-factor alpha.*

III. *There is a quadratic relationship between the number of exclusion and norms-based screening criteria and risk-adjusted return, when measured as the Sharpe ratio.*

Furthermore, the impact of individual screens, whether the fund engages in positive screening and more, are discussed but not extensively explored, as it is not the main research question.

1.4. Delimitations

Since there is far from any consensus on how externalities should be priced, conventional financial measures of performance are employed in the study. After deciding on the focus of the thesis, limitations mostly lie in the size of the data sample. The mutual funds examined are restricted to those included in "Hållbarhetsprofilen", which is an online database with information and guidelines concerning the asset selection process employed by managers of SRI mutual funds, the structure of the framework will be further discussed in the methodology section. Since information is not available on whether the mutual funds listed in "Hållbarhetsprofilen" applied the SRI practices before the introduction of the framework, a compromise had to be done between either assuming unchanged SRI approaches among the mutual funds to study a longer time period, or shortening the time frame in order to minimise the risk of changed SRI approaches over the years. The resulting sample period for measuring fund performance are the 36 months ranging from March 2011 – March 2014 and it is assumed that the funds' SRI approach remained the same, or at least similar enough not to affect inference.

Even though the ethical board of mutual fund marketing ("Etiska Nämnden för Fondmarknadsföring") perform audits and have stated guidelines of what is needed to market a fund as socially responsible, there is no strict regulation. Examining and quantifying the

screening process is a step further than simply classifying a fund as SRI or non-SRI, but it is difficult to know how seriously the policies presented to stakeholders actually are applied in the operational work. One could argue that this investigation somewhat becomes a study of the relationship between performance and *stated ambitions* of screening intensity.

To fully judge whether an investment is a good addition to an existing portfolio, covariance properties and the final exposure to systematic risk are important to consider. It is therefore difficult to make recommendations of an optimal mutual fund as the evaluation is dependent on the individual portfolio composition.

2. THEORY & LITERATURE REVIEW

The theoretical background for the methodology lies in portfolio theory and the pricing models the theory has brought forward as well as stakeholder theory because it appears as an explanatory theory in many previous SRI studies. The results of the earlier and most prominent papers are shortly summarised.

2.1. Portfolio Theory

Portfolio Theory, presented by Markowitz (1952), builds on the idea of combining assets to form a portfolio with the highest possible return given a certain amount of risk, or respectively, minimising the risk given the level of expected return. Risk is defined as standard deviation of returns and the theory relies on the mathematical proof that diversifying across assets lowers the overall risk compared to investing only in an individual asset. When the number of assets in the portfolio is large enough, dependent on their correlation characteristics, the idiosyncratic risk linked to a specific company is essentially eliminated and only the systematic risk affecting the global market remains.

When an additional asset is considered for a well-diversified portfolio, its covariance properties with the portfolio is of main focus as it increases or decreases the exposure to the systematic risk. Whether the specific risks of the assets are high or low is not of interest since the risk is diversified away and it is only how the asset are correlated that matters. It is assumed that investors are rational, non-satiated and risk-averse so that in choosing between portfolios higher return and lower risk is always preferred.

By estimating return, standard deviation and how assets correlate it is possible to weigh assets in a portfolio that minimises variance, this gives rise to a set of mean-variance efficient portfolios. If the investable universe is restricted, the number of assets that have favourable covariance properties with the portfolio will decrease and this could worsen the trade-off between expected return and standard deviation for the available set of portfolios. The reward of increasing risk with another unit of standard deviation can be measured with the Sharpe ratio, also referred to as reward-to-variability ratio.

2.1.1. Pricing models

An approach for studying whether an asset such as a mutual fund has performed worse or better than they ought to, given their risk factor sensitivities, is to compare actual returns with the theoretically fair returns derived from factor pricing models. Two measurements of this abnormal return are used in the thesis, Jensen's alpha from the Capital Asset Pricing Model (CAPM) and alpha from Carhart's four-factor model.

CAPM is an approach for estimating the appropriate discount rate used for pricing an asset's future cash flows. It builds upon the assumption that the portfolio is already well diversified, so for an asset with desirable covariance properties, the method to determine the appropriate level of expected return is to estimate the asset's β, the sensitivity to fluctuations in the market, in other words, the exposure to systematic risk. The reasoning is that holding a risky asset and bearing the risk of the market should be rewarded beyond the risk-free rate and the premium for this exposure is the excess return of the market, $[E(R_m) - R_f]$. With figures on the risk-free rate, the factor sensitivity and the return of the market it is then possible to statistically estimate the expected rate of return with equation (1).

$$E(R_i) = R_f + \beta_i[E(R_m) - r_f] \quad (1)$$

To allow for the possibility that there are more risk factors than market risk that is of importance when determining the appropriate level of expected return, Carhart (1997) built upon the work of Fama and French (1993) and developed a four-factor model. Factors considered are market risk, size, book-to-market ratio of companies and momentum of stock prices. The coefficients are estimated for each individual fund by performing an ordinary least squares regression with model (2).

$$R_t = \alpha_{carhart} + b(r_{m,t} - r_{f,t}) + s \times SMB_t + h \times HML_t + w \times WML_t + \varepsilon_t \quad (2)$$

Where R_t is excess return, $\alpha_{carhart}$ is the intercept, the abnormal return. The coefficients b, s, h and w are factor loadings for the corresponding risk factor premiums r_m-r_f, *SMB*, *HML* and *WML*. *SMB (Small Minus Big)* represents the performance of small stocks minus the performance of big stocks, *HML (High Minus Low)* is the difference in performance between

value and growth stocks as derived from their book-to-market ratio, *WML (Winners Minus Losers)* represents the momentum factor, it is the performance of winner stocks relative to loser stocks on an equally-weighted portfolio of small and big companies. The risk factor premiums are estimated by constructing factor-mimicking portfolios.

2.1.2. Efficient Market Hypothesis

An essential assumption of portfolio theory is that the market is efficient, meaning that there is no systematic mispricing of assets that can be profitably exploited on a regular basis. Fama (1970) defined three subsets of how security prices reflect all available information and adjust rapidly when new information is available. The three forms of market efficiency are called weak, semi-strong and strong form. In the weak form prices reflect all historical price- and volume information, in the semi-strong form also all publicly available information is reflected in prices and in the strong form even private, insider information is reflected in the market price. To allow for the possibility that superior analysis can lead to superior risk-adjusted returns on a regular basis after deducting the costs associated with gathering this additional information, the semi-strong (and the strong) form of the market efficiency hypothesis must be rejected.

2.2. Stakeholder Theory

In the Journal of Business Ethics, Miles (2012) stated the fundamental question of stakeholder theory as "who or what really counts" and that stakeholder management, holding other things equal, will enhance conventional corporate performance in terms of stability and growth. Stakeholders are defined as people or groups with sincere interest in a corporation's activities, e.g. shareholders, employees, customers, suppliers, governments, competitors and trade enhancing alliances. A further addition is that stakeholder theory can describe how cooperative and competitive benefits are connected to financial corporate goals and how these constellations of relationships possess value to the firm (Donaldson & Preston, 1995).

Michael Jensen (2002) stated "A firm cannot maximise value if it ignores the interest of its stakeholders". Active socially responsible fund managers try to influence the management of

firms in order to enhance the firms' ethical operations and these fund managers often use their voting rights and keep a dialogue with the firms' management since they most likely believe the time and effort put into these actions will increase the value of the investment.

Jensen (2002) questioned the recognised idea of value maximisation as well as stakeholder theory, when applied on their own, claiming that the former is among other things incapable of conveying a vision or purpose to participants in an organisation. The argument against practicing stakeholder theory on its own is that it does not sufficiently specify how to prioritise stakeholder interests in decision-making or how to assess performance. He further claimed that stakeholder theory "plays into the hands of special interests that wish to use the resources of corporations for their own ends". The suggested link between the two concepts is named "enlightened stakeholder theory", with one of the central principles being "We cannot maximise the long-term market value of an organisation if we ignore or mistreat any important constituency".

Socially responsible fund managers engaging with companies in their portfolio may believe the activities will create and uncover value and that recognising environmental, social and governance factors enhances the value maximisation potential of firms. If the fund manager is ahead of the rest of the market in this aspect, value can be found before the information is reflected in the price, a reasoning supported by Jensen; "I say "long-term" market value to recognize the possibility that financial markets, although forward looking, may not understand the full implications of a company's policies until they begin to show up in cash flows over time. In such cases, management must communicate to investors the policies' anticipated effect on value, and then wait for the market to catch up and recognize the real value of its decisions as reflected in increases in market share, customer and employee loyalty, and, finally, cash flows."

2.3. Literature Review

Much of previous research in the field has compared conventional fund performance to socially responsible fund performance to examine whether ethical investors pay a price for applying responsible policies to their investment strategies. The results among researchers differ and different methodological processes have been used in order to find a superior way of evaluating SRI mutual funds. Many researchers have investigated individual countries when evaluating SRI fund performance. Goldreyer & Diltz (1999) and Statman (2000) present results from mutual funds in the US and show that the difference between SRI and non-SRI mutual fund performance is not statistically different from zero. Moreover, Luther et al. (1992) Mallin et al. (1995) and Gregory et al. (1997) found similar results when they assessed SRI mutual fund performance in the UK. Bauer, Otten & Tourani Rad (2006) and Bauer, Derwall & Otten (2006) reached similar conclusions when evaluating SRI mutual fund performance in Canada and Australia. However, Geczy, Stanbaugh & Levin (2003) found results that the mean-variance optimisation was impaired by adding constraints when they examined socially responsible investments from the perspective of an investor who is only interested in US equity mutual funds with the highest Sharpe ratios.

Another approach that is frequently conducted related to SRI performance is multi-country analyses. Bauer et al. (2005) deeply studied US, UK and German mutual funds and Schröder (2004) performed studies on US, Swiss and German SRI mutual funds. Neither of these studies found any statistically differing results for conventional and SRI mutual fund performance. Renneboog, Ter Horst & Zhang (2008) contributed to the field of research by extensively studying the performance of nearly all SRI mutual funds in Europe, North America and Asia-Pacific over a longer time frame than most previous research on the topic and reached to the conclusion that France, Sweden, Ireland and Japan were the countries that significantly differed negatively from the performance of conventional mutual funds.

Barnett & Salomon (2006) published findings of a curvilinear, U-shaped relationship between the intensity of the socially responsible screening process and financial performance. They found a negative effect of additional negative screens but as more of these constraints are added to the portfolio, the relationship turns positive to almost reach the initial level of return for funds applying no screens at all. Their reasoning for the existence of a U-shaped

relationship was that mutual funds with low levels of screening intensity can still be well diversified and stay close to the mean-variance efficient frontier. Those mutual funds with high levels of screening intensity may suffer from not being fully diversified but are possibly compensated by the advantages of only investing in companies with high socially responsible standards, resulting in positive stakeholder relationships. Mutual funds that apply mid-level screening intensities will neither hold a mean-variant efficient portfolio nor a portfolio with enough investments in regard to high SRI quality.

The existence of a quadratic relationship is an occurrence that Lee, Humphrey, Benson & Ahn (2010) looked into as well in the article "Socially responsible investment fund performance: the impact on screening intensity". They did not observe any statistically significant results for a quadratic relationship with performance as Barnett & Salomon did, but they did however find a similar relationship when several different measures of risk were used as dependent variables.

As described, varying results have been found and different methodologies and data samples have been used. Most studies evaluate performance using the CAPM to determine whether there is a difference between conventional mutual fund performance to SRI mutual fund performance and plenty of research also employed Carhart's four-factor model by including the additional risk factors size, book-to-market and momentum.

3. DATA & METHODOLOGY

Despite drawing inspiration and guidance from esteemed papers in the area of research, the process of performing an empirical analysis is an ambiguous one and adjustments, assumptions and limitations have to be done along the way. To make this study possible to recreate for further developments, the data sources, approaches, issues and solutions are described starting with how the sample was selected, followed by the definition of screening intensity, explanations on how the various dependent variables of risk-adjusted return were estimated, which control variables are included and how the model for estimating the effect of screening intensity was defined and used.

In order to increase our understanding and especially forming a perception of how SRI is performed in practice, the research was initiated with an interview at the "Second AP Fund". Christina Olivecrona, responsible for sustainable investments and Jonas Eixmann, responsible for Swedish equities were of great help in improving our understanding of the topic, the focus of the study however lies in the quantitative analysis and the interview primarily served as an inspirational source.

3.1. Sample Selection

The starting point for gathering data on the different ways of approaching socially responsible investing in Sweden was the framework "Hållbarhetsprofilen". It is an initiative brought forward in 2011 by Swesif, the Swedish Sustainable Investment Forum, which is a sub-organisation to the international Eurosif. These organisations are independent, non-profit networks that promote socially responsible investments. "Hållbarhetsprofilen" provides an online database with information and guidelines concerning the asset selection process employed by managers of SRI mutual funds included in the Swedish Premium Pension System who have opted to publish a sustainability report along with the regular financial reporting. The focus of the SRI approach was placed mainly on the number of exclusion and norms-based screening criteria funds use, also referred to as negative screening. Eurosif reported in 2012 that almost 70 % of the Swedish SRI fund market is focused on this type of

strategy and it could be argued that the number of screens could be seen as a proxy for how extensively fund managers filter the investable universe in the asset selection process.

As of 2014-04-05, there was information on 110 mutual funds with varying investment style and geographical focus in "Hållbarhetsprofilen", for this study however, those that are not pure equity funds and those that have existed less than three years were excluded, resulting in 65 Swedish SRI mutual equity funds. It would be preferable to have an even larger sample size, and there are funds available in the market with names such as "ethical" or "sustainable" that are not listed in "Hållbarhetsprofilen". These funds could have been included, however, the fact that the funds in our sample report on their SRI-processes gives some verification that they truly act as socially responsible mutual funds. Also, unfortunately it would be too tedious and time-consuming for this study to contact each of these fund managers to inquire about their guidelines, if they would even be willing to share this information.

One could imagine that if a relationship does exist between the degree of screening intensity and fund performance, the strength of the relationship could differ over time as the economy fluctuates and as the possible value or cost of SRI is more or less clear to investors dependent on how heavily the topic is discussed in media. It would therefore be well advised to perform an analysis over a longer time frame to include more observed returns and adjust for any seasonality or trend. However, there is no data readily available as to whether the SRI approach of funds in the sample has differed during the funds' existence. Despite this lack of data, to enable a study of a long enough time period of returns, the assumption is made that funds' processes reported in "Hållbarhetsprofilen", since the introduction of its unified framework in 2011, have not changed to the extent that the changes would affect results and conclusions. Furthermore, SRI is a rapidly changing subject and this study primarily focuses on the role SRI plays in its current shape, during the period March 2011 – March 2014.

Another issue caused by the lack of historical data is survivorship bias. Brown, Goetzmann, Ibbotson & Ross (1992) analysed the relationship between risk and return and proposed that past fund performance foretells future performance. They uncovered patterns of predictability in performance and found significant evidence for the representative heuristic "hot hands", related to manager trackback and mutual fund performance. Although studies, among them Jensen (1968), demonstrate that future performance cannot be guaranteed, and that

outperforming the market on a risk-adjusted basis due to professional manager skill is rare, the fact remains that skilful managers may be those who survive. In line with Goetzmann and Ibbotson (1991), Hendricks, Patel, and Zeckhauser (1993) obtained strong results for that positive and negative performance can persist. Only including surviving funds in a sample is generally considered to result in upward bias as badly performing funds are those disappearing and the relatively worse performance is thus neglected in the analysis. However, one should be careful in what conclusions to draw about survivorship bias. Mutual funds that disappear may not necessarily have performed badly; funds with similar investment styles are often merged as a consequence of desiring a more specific and clear fund investment objective.

To investigate whether funds have disappeared from "Hållbarhetsprofilen" since its inception in 2011, an administrator at Swesif was contacted and unfortunately there are no such records available. Relying on previous studies by Renneboog et al. (2008) and Barnett & Salomon (2006) who tested for survivorship bias within SRI mutual funds and found that the attrition rate of SRI mutual funds was very low, the matter was not further explored in this study and we chose to examine the funds' performance over a three-year period to minimise the risk of survivorship bias.

3.2. Fund Managers' Screening Intensity

In contrast to the corresponding SRI-framework in the US, different types of screens being used by funds are not readily divided into "checkbox categories", but instead the SRI-processes are described freely under specified headlines. To make this qualitative data quantitative for analysis purposes, the information on processes was narrowed down to twelve distinct categories of negative screens; alcohol, tobacco, gambling, pornography, oil, military equipment and weapons, inhuman weapons (i.e. chemical, nuclear), corruption, pollution, child labour, violated labour rights and human rights. The number of screens a fund manager takes into account when excluding companies were noted for each fund. Six of the funds do not list any negative screening criteria and those were regarded as not engaging in negative screening. Although they simply use a different approach, for example restraining possible investments to only those selected for sustainable best-in-class indices, where one

could assume many negative screens are in fact taken into account, including the funds would distort the results of the applied method. The sample was thus reduced to 65 funds.

Many mutual funds do not only use negative screening criteria to exclude companies but also select companies based on that they actively work with trying to improve operations concerning environmental and social issues, and some follow a theme such as only investing in companies from an industry sector that is considered to be developing sustainability, i.e. clean energy. To account for the impact these activities could have on fund performance these two matters were noted as dummy variables for positive screening and theme investing respectively. Furthermore, a dummy was defined for whether the fund uses their proxy voting right on shareholders' meetings since this could be seen as an indication that fund managers try to generate change and engage with portfolio companies' board and management.

3.3. Fund performance

To calculate each funds monthly return for the 36 months ranging from 2011-03-31 to 2014-03-31, end-of-month closing prices for 37 months were gathered via Bloomberg and the month over month percentage return was calculated. In order to have comparable measures of return, the fact that the funds take on different levels of risk exposure is taken into account with three measures; Jensen's alpha (1968), alpha from Carhart's four-factor model (1997) and the Sharpe ratio (1966).

3.3.1. Jensen's alpha

When estimating coefficients of the CAPM for each of the funds, the intercept is the main parameter of interest as it measures the rate of return that either exceeds or falls short of what the "fair return" should have been when market risk is the only factor taken into account.

$$R_t = \alpha_{CAPM} + b\left(r_{m,t} - r_{f,t}\right) + \varepsilon_t \quad (3)$$

In the model R_t is excess return of the fund in month t, α_{CAPM} is the intercept and hence the Jensen's alpha, b is the factor loading, r_m is the return of the market and r_f the risk-free rate, ε_t

are the residuals from the fitted model. The model was estimated for each of the 65 funds in the sample by regressing excess return in the 36 months *t* on the excess market return from the market in the corresponding region. The historical factors of market risk and the risk-free rates were downloaded from Kenneth French's Data Library.

3.3.2. Carhart's four-factor alpha

To allow for the possibility that more than the exposure to systematic risk of the market is taken into account, a regression for each of the funds with Carhart's four-factor model is used that includes factors to account for size of assets, book-to-market ratio and momentum in prices.

$$R_t = \alpha_{carhart} + b\left(r_{m,t} - r_{f,t}\right) + s \times SMB_t + h \times HML_t + w \times WML_t + \varepsilon_t \qquad (4)$$

Where R_t is excess return, $\alpha_{carhart}$ is the intercept, the abnormal return, i.e. the parameter of interest. The coefficients b, s, h, and w are factor loadings for the corresponding risk factors of market exposure (r_m-r_f), size (*SMB)*, book-to-market ratio (*HML)* and momentum (*WML)*.

The Kenneth French Data Library provides estimated factor premiums for markets in Asia Pacific excluding Japan, Europe, Japan, and North America as well as for a global portfolio. We assigned each fund to one of these regions based on where the majority of their investments are held and then regressed on the matching set of factors to increase the accuracy of the factor loading estimations. The consistency gained from gathering data on factors from the same source is advantageous for inference purposes as the data were estimated with the same methodological process.

As both of these regression models' dependent variable is the excess return of the fund, the US 1-month Treasury bill rate from the corresponding month was subtracted from each of the previously calculated returns. The reason for using the US T-bill rate is that it is the one applied by Fama and French when estimating the risk premiums and using the same rate provides uniformity. Furthermore, with the notion of global financial markets, the US T-bill rate is often considered as the risk-free benchmark rate (Ma, Tchen, Smith & MacNamara, 2011).

Both of the regressions (3) and (4), with the resulting estimations of intercepts to be used as the dependent variables later on in main model (6), were estimated with the Ordinary Least Squares (OLS) method. Many of the estimates for both Jensen's alpha and Carhart's four-factor alpha are insignificant when examined with a standard t-test, possibly a consequence of the time frame limitation in sample size, along with the randomness often observed in asset prices. A way of addressing this problem was presented by Chan & Faff (2003) as they recorded the standard errors of these estimates and used them as weights in a Weighted Least Squares (WLS) regression in the main model investigating the impact of screening intensity. This approach was tested but as it changed the inferences on several points, since it is recommended to avoid setting weights if uncertain, and as it was not employed by neither Barnett & Salomon (2006) nor Renneboog et al. (2008), the OLS method was relied on and the WLS results will not be reported.

3.3.3. Sharpe Ratio

When discussing factor-pricing models it is often assumed that only systematic, or factor risks are rewarded in the form of risk premiums. However, Malkiel and Xu (1997) argued that idiosyncratic risk is also rewarded. To allow for the possibility that idiosyncratic risk actually do have an impact on performance we include the funds' Sharpe ratios (1966) as another measurement of performance as it takes standard deviation into account and therefore is a measurement of performance relative to total risk.

$$SR_i = \frac{\overline{r_i} - \overline{r_f}}{\sigma_i} \qquad (5)$$

Here $r_i - r_f$ is the average excess return over the 36 months and σ_i is the standard deviation of the excess returns. The Sharpe ratio is a measure of the excess return rewarded for each additional unit of standard deviation and is widely used when evaluating mutual fund performance. Since the Sharpe ratio is equal to the slope of the capital allocation line where investors theoretically choose between the risk-free asset and the optimal risky portfolio it is a good measure for examining how negative screens affect the risk-reward characteristics of the optimal risky portfolio.

3.4. Control Variables

Three control variables for fund characteristics are included in main model (6) to account for the impact that age, size and fee might have on performance. Age is measured in months since the funds' inception date up until April 2014, it accounts for the plausible learning effect of fund managers. Size is the total assets under management measured in millions of SEK as of 2014-03-31. The variable is included due to the advantages and disadvantages linked to managing large amounts of assets. Fee is the stated annual management fee in percentages and retrieved 2014-04-03.

When downloading data on control variables some values were not available from Bloomberg and those were instead gathered via Morningstar. After noting some discrepancies between these two sources concerning fund inception date and noting that Morningstar was the one in line with the funds' webpage data, Morningstar was solely relied on to provide data on the funds' inception dates.

3.5. Parameter Estimation

For the purpose of investigating how the number of screens employed by portfolio managers of SRI funds relate to risk-adjusted performance, the approach used is similar to the one first introduced by Barnett & Salomon (2006). This approach was further developed by Renneboog et al. (2008) as they assessed a study of SRI-fund performance with a wider geographical focus, and later D.D. Lee et al. (2010) proceeded with a study building upon the work of Renneboog et al. (2008). To better suit our sample characteristics and data availability, minor adjustments in the model and a cross-sectional approach was used.

$$\begin{aligned} Performance_i = {} & \beta_0 + \beta_1 \times inty_i + \beta_2 \times inty_i^2 + \beta_3 \times theme_i + \beta_4 \times posscr_i \\ & + \beta_5 \times vote_i + \beta_6 \times age_i + \beta_7 \times size_i + \beta_8 \times fee_i + \beta_9 \times asiapac_i \\ & + \beta_{10} \times europe_i + \beta_{11} \times japan_i + \beta_{12} \times northamerica_i + \varepsilon_i \end{aligned} \quad (6)$$

$Performance_i$ is the dependent variable in the model with a value for each fund i, represented in three different regression outputs by the measures Jensen's alpha, Carhart's four-factor alpha and the Sharpe ratio. The main variable of interest, which the performance measures

will be regressed on, is the independent variable $inty_i$ and the variable represents the number of negative screens listed by each fund in "Hållbarhetsprofilen". The possibility of a quadratic linear relationship was taken into account with the variable $inty^2{}_i$

Furthermore, dummies related to the mutual funds' SRI approach are $posscr_i$, equal to 1 if the fund uses positive screens to select socially responsible companies and 0 otherwise, $theme_i$, set to 1 if the fund has a theme investment strategy. The dummy $vote_i$ has a value of 1 if the asset management company attends shareholder meetings and use their voting rights.

The factors age_i, $size_i$ and fee_i are control variables and $asiapac_i$, $europe_i$, $japan_i$ and $northamerica_i$ are dummy control variables because the funds invest in different regions. These dummies are included to account for regional differences in performance beyond that derived from the risk premiums. To prevent perfect collinearity the global mutual funds serve as base group, which the other mutual funds are compared to.

3.5.1. Individual screens

To investigate whether certain negative screens have a significant impact on performance on their own, the three measures of performance were also regressed on all of the negative screens in the form of individual dummies, set to 1 if the fund employs the screen and 0 otherwise.

$$\begin{aligned} Performance_i = {} & \beta_0 + \sum_{j=1}^{12} \beta_j \times screen_{i,j} + \beta_{13} \times theme_i + \beta_{14} \times posscr_i \\ & + \beta_{15} \times vote_i + \beta_{16} \times age_i + \beta_{17} \times size_i + \beta_{18} \times fee_i \\ & + \beta_{19} \times asiapac_i + \beta_{20} \times europe_i + \beta_{21} \times japan_i \\ & + \beta_{22} \times northamerica_i + \varepsilon_i \end{aligned} \qquad (7)$$

The twelve different screens appear with their corresponding coefficients β_j first in the regression and the remaining variables are the same as in model (6).

3.5.2. Heteroskedasticity

The White-test showed no signs of heteroskedasticity in the regression models for any of the three performance measures used as dependent variables. A Breusch-Pagan test however rejected the null hypothesis of homoscedasticity at the 5 % significance level for all three measures and thus indicated that the residuals in the model might be correlated with the independent variables. Due to this ambiguity regarding heteroskedasticity robust standard errors were calculated. The Eicker-Huber-White standard errors are only valid for large samples, but instead of judging whether the sample is large enough they are simply presented together with the regular standard errors.

As for regressing on the individual screens, the Breusch-Pagan test showed signs of heteroskedasticity for all the performance measures at various significance levels, whereas the simple White-test pointed to heteroskedasticity only for Carhart's four-factor alpha and not for Jensen's alpha or for the Sharpe ratio. Therefore the same approach of reporting both regular and robust standard errors was done.

3.6. Summary Statistics

The sample consists of 65 Swedish mutual fund observations and the summary statistics of all variables are presented in Table 3.

Table 3: Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Jensen's Alpha	65	0.1257	0.4096	-1.5444	1.0939
Carhart's Alpha	65	-0.0626	0.3995	-1.3119	1.1538
Sharpe Ratio	65	0.1783	0.1270	-0.1320	0.4911
Screening Intensity	65	8	2.1139	4	11
Theme	65	0.0615	0.2421	0	1
Positive Screening	65	0.3692	0.4863	0	1
Vote	65	0.3538	0.4818	0	1
Age	65	172.6154	75.7833	21	408
Size	65	5450.289	21573.09	72.44	173497
Fee	65	1.0127	0.5375	0.15	2
Global	65	0.3538	0.4818	0	1
Europa	65	0.4615	0.5023	0	1
Japan	65	0.0461	0.2114	0	1
Asia Pacific	65	0.0769	0.2685	0	1
North America	65	0.0615	0.2421	0	1

Jensen’s alpha and Carhart’s four-factor alpha are measured monthly in percentages. A Jensen’s alpha of one means that the mutual fund has on average returned a one percentage point higher return than what it was expected to, given its level of exposure to market risk by correlating with a wide market index in its respective region. The Sharpe ratio is the return gained per unit of standard deviation of returns, for example a difference in Sharpe ratio of 0.2 between mutual funds would mean that the superior fund generated 0.2 percentage points higher return for each percentage point of monthly standard deviation the fund has had on average over the last three years.

4. RESULTS

The results of regressing performance on screening intensity are presented and followed by explanations of the economical and statistical significance. To serve as additional analysis, the performance measures are regressed on each of the twelve individual screening criteria. Finally the hypotheses stated in the introduction are answered.

4.1. Screening Intensity

The resulting outputs from regressing performance on screening intensity as specified in model (6) are presented for each of the examined measures Jensen's alpha, Carhart's four-factor alpha and the Sharpe ratio along with the regular and robust standard errors in Table 4.

Table 4: Exclusion and Norms-Based Screening Intensity			
Variable	Jensen's alpha	Carhart's alpha	Sharpe Ratio
Screening intensity	0.8222 (0.1941)*** [0.2430]***	0.2524 (0.1944) [0.2732]	0.1316 (0.0530)** [0.0557]**
Screening int. squared	-0.0481 (0.0119)*** [0.0147]***	-0.0159 (0.0120) [0.0161]	-0.0077 (0.0032)** [0.0034]**
Theme investing	0.1650 (0.1805) [0.1568]	0.0716 (0.1807) [0.1490]	0.0332 (0.0492) [0.0486]
Positive screening	-0.1689 (0.0899)* [0.0909]*	-0.1221 (0.0900) [0.0843]	-0.0297 (0.0245) [0.0249]
Vote	0.0943 (0.0900) [0.0917]	-0.0204 (0.0901) [0.0897]	0.0053 (0.0245) [0.0260]
Age	0.0007 (0.0006) [0.0006]	0.0003 (0.0006) [0.0005]	0.0002 (0.0001) [0.0002]
Size	2.13e-07 (2.00e-06) [8.04e-07]	-4.72e-07 (2.00e-06) [6.87e-07]	1.66e-07 (5.45e-07) [2.34e-07]
Fee	-0.0476 (0.0826) [0.0776]	-0.0083 (0.0827) [0.0809]	-0.0267 (0.0225) [0.0222]
Europe	-0.1082 (0.0970) [0.1048]	-0.4794 (0.0971)*** [0.0992]***	-0.0800 (0.0264)*** [0.0297]***
Japan	-0.5242 (0.2023)** [0.1101]***	-0.5153 (0.2026)** [0.1163]***	-0.2504 (0.0552)*** [0.0302]***
Asia Pacific	-0.3112 (0.1630)* [0.1259]**	-0.3190 (0.1632)* [0.1162]***	-0.1979 (0.0445)*** [0.0381]***
North America	0.4386 (0.1793)** [0.1107]***	0.3056 (0.1795)* [0.1162]**	0.1736 (0.0489)*** [0.0383]***
Constant	-3.1441 (0.7551)*** [0.9497]***	-0.7461 (0.7561) [1.0763]	-0.3022 (0.2061) [0.2061]
R^2	0.5232	0.4974	0.6306

Results from equation (4). Standard error in ( ), robust standard errors in [ ].
Statistically significant at: *10 %, **5 %, *** 1 %

4.1.1. Jensen's alpha

The variables of highest relevance for inference, screening intensity and screening intensity squared, provide statistically significant results when Jensen's alpha is the performance measure and thus the dependent variable. Screening intensity has a positive coefficient of 0.8222 and the output is statistically significant with both regular and robust standard errors at the 1% level, which is a strong indication of accurate results. The value of 0.8222 means that one additional screen is correlated with an increase in alpha of 0.82 percentage points in monthly average return over the sample period. This may seem to be of large economic value as it is possible to apply up to 12 different screens, however, the effect wears off due to the existence of a curvilinear relationship represented in the model by the squared counterpart. The variable screening intensity squared is statistically significant with regular and robust standard errors at the 1% level as well. The relatively smaller coefficient of -0.0481 indicates that the positive effect of screening intensity diminishes and as the number of screens increases the negative coefficient from the squared variable starts to have a significant economic impact on alpha. The main focus for analytical purposes however lies in the joint effect of these two variables, as there is a point where an additional screen no longer enhances portfolio performance.

To exemplify the joint impact, applying four screens would imply $4 * 0.8222 + 4^2 * (-0.0481) = 2{,}5192$ percentage points higher monthly alpha on average compared to employing zero screening criteria, whereas ten screens correlate with $10 * 0.8222 + 10^2 * (-0.0481) = 3{,}412$ percentage points higher alpha. To employ nine screens is linked to 3,5037 percentage points, which is a higher alpha than when ten screens are applied to the investment strategy. These calculations show that there is an optimal number of screens for maximising the fitted value of the model before the squared effect is large enough that refraining from applying another screen is favourable. In line with the interpretation of Figure 1, calculating the derivative shows that the turning point lies at 8.55, in other words at nine negative screens.

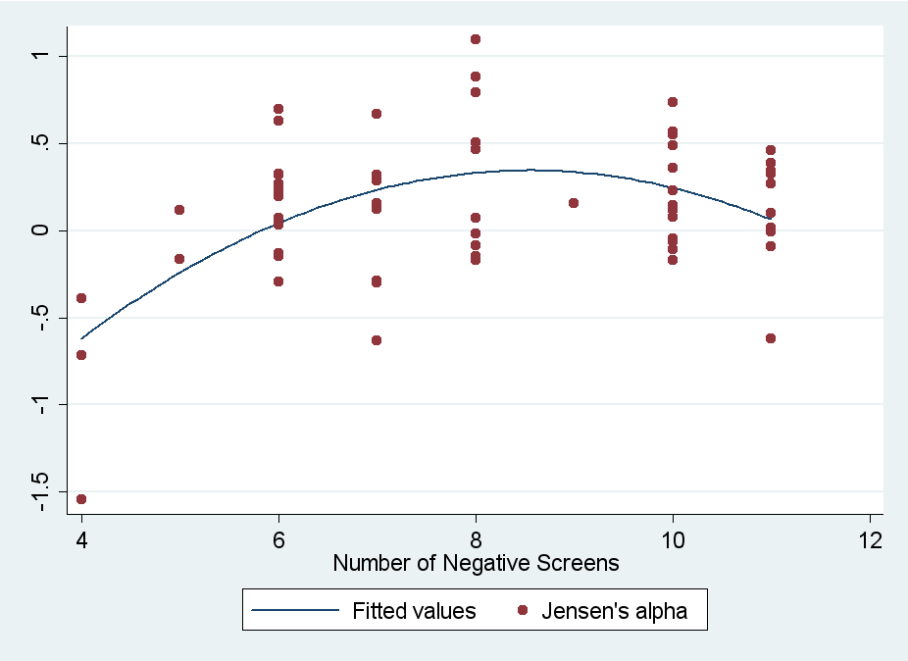

Figure 1: Quadratic Relationship of Screening Intensity

Figure 1 showing the quadratic relationship of screening intensity demonstrates the curvilinear relationship between abnormal monthly return and the number of exclusion and norms-based screens. To test the necessity of a quadratic relationship with screening intensity, a regression was performed without the squared variable. Without including the squared screening intensity in the regression, the coefficient of screening intensity is 0.0466 and statistically significant at the 10% level with robust standard errors, and at the 5% level with regular standard errors. By including the squared variable, the model allows for a curved relationship instead of a linear one, R^2 measuring goodness of fit increases and the observed result are more statistically significant.

The dummy variable positive screening has a negative coefficient of -0.1689 and is statistically significant with regular and robust standard errors at the 10% level. The negative coefficient signifies a shift downwards in alpha of 0.17 percentage points. This means that the funds employing positive screening has on average, during our sample period, delivered 0.17 percentage points lower abnormal return each month, given its exposure to the market, compared to the funds that do not use positive screening to select socially responsible assets.

The regional dummy variables for Europe, Japan, Asia Pacific and North America are control variables to account for the regional differences in risk-adjusted performance these markets have had during the sample period. The coefficients for the regional variables are compared

to the global portfolios that serve as base group. Japan has statistically significant results at the 1 % significance level with regular standard errors and at 5% with robust standard errors and the region has a negative coefficient of -0.5242, meaning on average, 0.52 percentage points lower alpha each month compared to the global portfolio. Asia Pacific also demonstrates a negative relationship with a coefficient of -0.3112 and statistically significant at the 10% level with regular and robust standard errors at the 5% level, which implies a 0.13 percentage points lower alpha than the abnormal returns of the global portfolios in the sample. North America is shown to have a statistically significant shift upwards in alpha of 0.44 percentage points, as the coefficient has the positive value 0.4386, statistically significant with regular standard errors at the 5% level and with robust standard errors at the 1% level.

4.1.2. Carhart's four-factor alpha

When Carhart's four-factor alpha is regressed on the same variables as Jensen's alpha, the relationship between screening intensity and abnormal performance is not statistically significant. The coefficient of screening intensity is 0.2524, which can be considered to be of economic significance, the squared counterparts' coefficient is -0.0159. These coefficients are however not statistically significant unless one trusts a 20 % significance level, which is not considered accurate enough to draw inference from.

The dummy variables of the geographic regions show some significant results. Europe is statistically significant at the 1% level with a coefficient of -0.4794, representing a lower average monthly alpha of funds in Europe during the period March 2011 – March 2014 compared to the global funds. Likewise, Japan had a negative coefficient of -0.5153 and the output is statistically significant at 5 % with regular standard errors and at 1 % with robust standard errors. Asia Pacific shows a coefficient of -0.3190, significant with regular standard errors at 10 % and with robust standard errors at the 1% level. Funds investing in North America were the only funds outperforming the global funds as the region had a higher alpha of 0.31 percentage points each month on average, significant at the 5 % level with robust standard errors and at 10 % with regular standard errors. However, these variables primarily serve as control variables, they are essentially a comparison of the average risk-adjusted

outperformance in these regions during the sample period and since they are not interacted with screening intensity they do not provide inference of much interest.

4.1.3. Sharpe Ratio

The quadratic relationship observed for Jensen's alpha is also present for the Sharpe ratio, showing that additional negative screens increase the ratio of return to standard deviation, at a decreasing rate, and at a certain point the effect of another screen is no longer advantageous. An increase in the Sharpe ratio is always preferred as it implies that the returns are either higher, the standard deviation lower, or both.

The relationship between performance and screening intensity is positive with a coefficient of 0.1316 and statistically significant with regular and robust standard errors at the 5 % level. This is to be interpreted such that one more screen is highly correlated with a 0.13 percentage points higher Sharpe ratio. This translates into the average monthly return for each percentage point of standard deviation to be 0.13 percentage points higher. In line with the outputs from the Jensen's alpha regression, screening intensity squared has a negative coefficient. It is significant at the 5 % level with regular and robust standard errors and has the economic value of -0.0077. This validates the negative impact that another screen starts to have on abnormal performance after a turning point, calculated with a derivation to occur at 8.55 screens, in line with the results found with Jensen's alpha. Applying nine exclusion and norms-based screening criteria would statistically have implied a $9 * 0.1316 + 9^2 * (-0.0077) = 0.5607$ higher Sharpe ratio during the sample period, a notable difference, compared to those funds applying zero screens when keeping all else constant.

As for the regional differences, all of them are statistically significant at the 1 % level, with both robust and regular standard errors. Europe, Japan and Asia Pacific all have negative coefficients of -0.08, -0.2504 and -0.1979 respectively whereas North America has a positive coefficient of 0.1736.

In contrast to Jensen's alpha, positive screening does not have a statistically significant impact on the Sharpe ratio, neither does theme investing nor the control variables age, size or fee for any of the three performance measurements.

4.2. Individual Screens

As the inference drawn from regressing Jensen's alpha and the Sharpe ratio on screening intensity showed that nine exclusion and norms-based criteria is the optimal number of screens, it is of course interesting to know what nine screens are the ideal ones. To serve as additional analysis in this matter, the performance measures were regressed on each of the twelve individual exclusion and norms-based screening criteria. Their individual impacts on the three performance measures are presented in Table 5. In the regressions, the variable pollution was omitted to prevent the issue of constant variables as all of the funds employ the screen, and the control variables are not reported for the sake of brevity. Some significant results were found and the reasons to why the observed relationship exist can be contemplated, but in order to build a thorough conclusion in regards to the optimal set of nine screens, further research is needed as consideration must be given to covariance properties, individual preferences, variations in regions and more.

Table 5: Individual Negative Screens			
Variable	Jensen's alpha	Carhart's alpha	Sharpe ratio
Alcohol	0.4714 (0.7869) [0.5687]	0.0345 (0.7606) [0.5032]	0.1767 (0.2078) [0.1590]
Tobacco	0.6529 (0.6227) [0.3898]	0.3860 (0.6019) [0.3619]	0.2372 (0.1644) [0.1184]*
Gambling	0.0493 (0.1705) [0.1988]	-0.0647 (0.1648) [0.1670]	0.0020 (0.0450) [0.0530]
Pornography	0.6488 (0.1714)*** [0.2339]***	0.5077 (0.1656)*** [0.2374]**	0.1383 (0.0452)*** [0.0621]**
Oil	-0.4976 (0.3086) [0.3094]	-0.2145 (0.2983) [0.3138]	-0.0823 (0.0815) [0.0837]
Weapons & Equipment	-0.6828 (0.7945) [0.6163]	-0.2653 (0.7679) [0.5581]	-0.2176 (0.2098) [0.1703]
Inhuman Weapons	0.2238 (0.2084) [0.2779]	0.2497 (0.2014) [0.2521]	0.0657 (0.0550) [0.0755]
Corruption	-0.5837 (0.4309) [0.2549]**	-0.2007 (0.4165) [0.2721]	-0.1794 (0.1138) [0.0709]**
Pollution	(Omitted)	(Omitted)	(Omitted)
Child Labour	-0.2767 (0.4829) [0.2683]	-0.0424 (0.4667) [0.1768]	-0.1254 (0.1275) [0.0767]
Violated Labour Rights	0.0988 (0.3312) [0.1173]	-0.1487 (0.3201) [0.1153]	0.0502 (0.0874) [0.0359]
Violated Human Rights	0.5195 (0.6600) [0.5529]	-0.2697 (0.6379) [0.4948]	0.1421 (0.1743) [0.1496]
R^2	0.6034	0.6105	0.7124

Results from equation (5). Standard error in ( ), robust standard errors in [ ].
Statistically significant at: *10 %, **5 %, *** 1 %

The regression outputs show that excluding companies involved with revenues from the pornography industry has a positive impact and is statistically significant when regressed with Jensen's alpha, Carhart's alpha and the Sharpe ratio as dependent variables. The screen is statistically different from zero with regular and robust standard errors at the 1% level and

has a positive coefficient of 0.6488 when Jensen's alpha is the measure of performance. This is an indication that funds excluding companies associated with the pornography industry on average had 0.65 percentage points higher abnormal monthly returns compared to those who did not exclude based on the screen pornography. The screen has a coefficient of 0.5077 when regressed with Carhart's alpha and is statistically significant with regular standard errors at 1% and at 5% with robust standard errors. As for the Sharpe ratio, the coefficient of 0.1383 is also statistically significant with regular and robust standard errors at the 1% level and 5% level respectively. It seems unlikely however that the mutual funds would have a big enough exposure to those, apparently underperforming companies that it would have such a notable positive impact to exclude them. One should perhaps not put too much weight on the inference, it certainly is statistically significant but there could be a correlation with some other, unknown factor.

Another individual screen that shows significant results is corruption. By excluding corrupt companies the abnormal performance of those socially responsible investment funds in the sample decreased. These results are statistically significant with robust standard errors at the 5% level with Jensen's alpha and the Sharpe ratio, with corresponding coefficients of -0.537 and -0.1794.

The abnormal performance was on average enhanced with a coefficient of 0.2372 when measured with the Sharpe ratio for those funds excluding tobacco companies from their investment strategy. The results are statistically significant with robust standard errors at the 10% level.

Even though most of the screens do not have a significant impact with a t-test, it is possible that they could have a joint significance and a F-test can show whether this is the case. Two F-tests were performed, one test for the exclusion screens alcohol, tobacco, gambling, oil, weapons and inhuman weapons (pornography excluded as the screen is significant on its own), and one test for the norms-based screens corruption, pollution, child labour, violated labour rights and violated human rights. The exclusion screening test proved significant at the 10 % level for Jensen's alpha and the norms-based screening test was not significant. As for Carhart's alpha and the Sharpe ratio, neither of the F-tests were significant.

4.3. Answering Hypotheses

The hypotheses to answer are whether a statistically significant quadratic relationship exists between the number of negative screens and the three measurements of performance. The approach of supporting the assumption of an existing relationship is to reject the null hypothesis stating that a relationship does not exist, in other words, to show that the coefficients are different from zero.

The regression model (6) is of interest, where performance was regressed on screening intensity and screening intensity squared in addition to the supplementary variables that are not relevant for hypothesis testing. The coefficient in model (6) for screening intensity is β_1 and the squared counterpart is represented by β_2. The null hypothesis states that none of the variables have an effect on the corresponding performance measurements and is consequently specified as follows.

$$H_0: \beta_1 = \beta_2 = 0 \quad (7)$$

The alternative hypothesis states that the variables do have an effect on performance and is thus expressed as both screening intensity and screening intensity squared being different from zero. Note that both variables must be different from zero to allow for a quadratic effect; if only the linear screening intensity is significant the null hypothesis cannot be rejected.

$$H_1: \beta_1 \neq 0 \; and \; \beta_2 \neq 0 \quad (8)$$

When performing a F-test, the goodness-of-fit of the original model is compared to the goodness-of-fit when the variables of interest are removed from the regression model, in this case screening intensity and screening intensity squared. The t-tests performed when determining whether the variables screening intensity and screening intensity squared were significant on their own show that the null hypothesis can be rejected for both Jensen's alpha and the Sharpe ratio and it is thus not necessary to examine the matter further with a F-test, as both of the variables were highly significant. However, it is of interest to assess at what level the null hypothesis of their joint impact can be rejected for Jensen's alpha and the Sharpe

ratio as well as testing whether the two variables joint effect might be important for a well-fitting model when Carhart's alpha is the dependent variable, even though the individual t-tests showed no significant relationship. The F-tests were done for each of the performance measurements – hypothesis I, II & III respectively, and returned a p-value, the probability of observing a value as extreme as we did if the null hypothesis is in fact true, which is the basis for our inferences.

I. Jensen's alpha - The null hypothesis is rejected at the 1 % significance level.

The F-test shows with high confidence that screening intensity had a quadratic impact on Jensen's alpha during the sample period. The chance of faulty rejecting the null hypothesis is less than 1 % and it is thus assumed that screening intensity has explanatory power and that the relationship is quadratic.

II. Carhart's alpha - The null hypothesis is not rejected.

For Carhart's alpha there is no clear evidence that screening intensity is of importance for modelling what has an effect on the abnormal return after the risk exposure to the market, company size, company book-to-market ratio and momentum has been accounted for.

III. Sharpe ratio - The null hypothesis is rejected at the 5 % significance level.

When the Sharpe ratio is the measure of performance and screening intensity and its squared counterpart are included as variables, the model's goodness-of-fit and explanatory power is improved compared to when they are not included. There is less than 5 % chance of wrongfully stating that screening intensity's effect on the Sharpe ratio is different from zero and that the relationship is quadratic.

5. DISCUSSION

The results retrieved when using Jensen's alpha and the Sharpe ratio as dependent variables indicate that portfolio managers aiming to achieve superior fund performance should ideally apply nine out of twelve applicable screening criteria. This is contradictory to the study by Barnett & Salomon (2006) in which a convex relationship was found, whereas regressions in this study resulted in a concave relationship. Their analysis stated that the funds are best off by either applying few screens, thus staying well diversified, or by applying many screens to truly gain the benefits derived from positive stakeholder relations.

Reasons to why the relationship found in this study might differ from previous research can have numerous explanations. As the quantitative measurements of SRI differ from previous studies, the results are not completely comparable. The approach of using screening intensity is similar, however the screens that fund managers in different countries employ vary due to the demands of retail investors stemming from differences in culture and public opinion. In the United States for example, screens include animal testing, excessive pay, country exclusions and more. Among the Swedish mutual funds in "Hållbarhetsprofilen" there are no portfolio managers stating that they apply any of these constraints.

Previous studies presented have had sample periods ranging from the 1980's up until the early 21st century whereas this study covers the last three years, a shorter time period but more up to date. As ideals and opinions towards sustainability fluctuate over time, the companies SRI mutual funds invest in may experience time periods in which the public opinion is beneficial and more value is put on the Stakeholder advantages than at other times when the issue is not given as much attention.

What separates Carhart's four-factor alpha from Jensen's alpha is that the additional factors of company size, book-to-market ratio and momentum have been taken into account when estimating abnormal return. The statistically significant relationship between abnormal return and screening intensity found for Jensen's alpha is lost when the estimation model includes the additional factor premiums. In March 2011 – March 2014 screening intensity did not offer any assistance in seeking to discover superior performance when Carhart's four-factor alpha is the measurement of performance. Since the abnormal return's correlation with

screening intensity is not statistically significant once the three additional factors are considered, it is imaginable that the exclusion and norms-based screening criteria are correlated with one or more of those factors. Applying more screens is linked to a higher alpha when using the CAPM, but the reason might be that more screens actually lead the portfolio manager to select companies with smaller market caps, higher book-to-market ratios or companies with positive momentum and that it is in fact biases from those factors, and not some analytical edge, that generate the superior Jensen's alpha.

As for Jensen's alpha and the Sharpe ratio, the null hypothesis that screening intensity and screening intensity squared do not have an impact, was rejected at the 1 % and the 5 % significance level respectively. According to portfolio theory, the funds with high levels of screening intensity, excluding many possible investment opportunities, should suffer from restricting the investable universe and not being fully diversified, thus deteriorating the chances of successful portfolio management. Mutual funds applying the least amount of constraints to their investment strategy should be better off at composing an optimal portfolio at the efficient frontier. By adding constraints, the minimum variance frontier will shift to a less optimal trade-off between risk and return and it is therefore according to portfolio theory suggested not to add more constraints to your portfolio than necessary.

That more screens enhance performance up to the turning point at nine screens, even though portfolio theory discourages restrictions, could be explained in the light of stakeholder theory. A more intense screening process when composing the portfolio will lead to investing in more healthy companies that are in accordance with social norms and stakeholder values. If emotions do have an impact on decision-making as shown by behavioural economists, one could imagine that sin companies, which are corporations associated with exclusion criteria, are companies that could easily be the subject of strong negative emotions, consequently reducing market value as some investors do not wish to be associated with such businesses.

Supposing that markets might not be fully efficient in its semi-strong form opens up for the possibility of profitable investment opportunities gained from superior asset analysis. Olivecrona and Eixmann at the Second AP Fund regard ESG analysis primarily as a tool for identifying the most cost efficient companies rather than solely as a tool for excluding sinful companies. Even though environmental, social and governance (ESG) considerations have

become common factors to consider in equity research, it is not a wild suggestion that SRI mutual funds incorporate ESG analysis to a wider extent than conventional mutual funds. Mutual funds with a higher number of screening criteria might have to be more scrutinising in the asset selection process, for analysts to discover whether a company respects labour rights, corruption laws or other norms-based criteria, they may have to perform a more rigorous analysis than to solely analysing financial data.

As the positive effects of screening starts to decline after nine screens, one can reason that the positive effects derived from stakeholder theory starts to diminish and the negative effects derived from having a constringent pool of possible investments starts to dominate.

There are some apparent contradictions in combining portfolio theory and stakeholder theory. The latter emphasises the possible gains to be derived from investing in companies with positive relations with stakeholders as the intrinsic value can be enhanced through constant improvements like the actions presented in Table 1, actions that are largely aimed at reducing company specific risks. Portfolio Theory primarily focuses on the covariance properties, as the company specific risks are diversified away and are therefore not relevant. However, the possible occasion for combining these two theories could be in that the covariance properties portfolio theory bases its assessments on are *estimates* of return, standard deviation and correlations and there are growing evidence that a detailed analysis incorporating stakeholder theory and ESG analysis can actually make those estimates more realistic and precise (Goldman Sachs, 2007). In the fast-paced and highly competitive financial markets unforeseen events are to be expected and it is likely that the covariance properties change. Even if company-specific risk is not of interest, it is reasonable to assume that the high-risk companies' estimates suffer a higher probability of changing due to unforeseen events, for example litigation, reputational risk or conflicts of interest, which is where stakeholder theory can be used in proactive purposes.

6. CONCLUSION & OUTLOOK

Studying and quantifying the common SRI practices led to the approach of reckoning categories of SRI-constraints, i.e. screens. From studying the performance of SRI mutual funds on the Swedish market over a three-year period a curvilinear relationship between risk and return was found. The regression analysis show that applying nine exclusion and norms-based screens would have been the optimal screening intensity for an investor choosing a SRI mutual fund on the Swedish market in March 2011 – March 2014.

For further research it would be interesting to quantify the mutual funds' characteristics more thoroughly by evaluating the quality of their assessments and deeply examine whether fancy words are primarily a marketing trick or if the words are truly incorporated in the organisation. It is worth to bear in mind that social responsibility is a highly individual subject and what branches of responsible investing are the most valuable differ among people, which makes evaluating the plausible value of different SRI practices hard to assess. In regards to exclusion and norms-based screening, light has been shed on the optimal screening intensity, and a further analysis on individual criteria would be interesting for future research.

BIBLIOGRAPHY

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

Bauer, R. & Smeets, P. (2013). Social Identification and Investment Decisions. Retrieved from: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2140856>. [18 April 2014].

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

Bauer, R., Otten, R., Tourani Rad, A., 2006. Ethical investing in Australia: is there a financial penalty? *Pacific-Basin Finance Journal* 14 (1), 33–48.

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

Bollen, N.P. (2007). Mutual fund attributes and investor behaviour, *Journal of Financial and Quantitative Analysis,* 42, 915-937.

Brown, S.J., Goetzmann, W., Ibbotson, R.G. & Ross S.A. (1992). Survivorship Bias in Performance Studies, *The Review of Financial Studies,* 5 (4), 553-580.

Capelle Blancard, G. & Monjon, S. (2012). The Performance of Socially Responsible Funds: Does the Screening Process Matter? *European financial management*. Retrieved from: <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-036X.2012.00643.x/abstract >. [3 March 2014]

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

Chan H.W. & Faff R.W. (2003). An investigation into the role of liquidity in asset pricing: Australian evidence, *Pacific-Basin Finance Journal,* 11 (5), 555-572.

Donaldson T. & Preston L.E. (1995). The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. *The Academy of Management Review*. 20 (1), 65-91.

Eurosif (2012). European SRI Study 2012 | Sweden. Retrieved from: <http://www.eurosif.org/research/eurosif-sri-study/sri-study-2012>. [13 May 2014].

ESG Managers® (2014), What is ESG?, Retrieved from: <http://www.esgmanagers.com/Sustainable_Investing/What_is_ESG>. [13 May 2014].

Fama, E.F. (1970), Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance* 25 (2), 383-417.

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

Fama and French 2014, *Data library – Developed market factors and returns*. Available from: <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>. [6 April 2014].

Fehr, E. & Fischbacher, U. (2003). The nature of human altruism, *Nature*, 425, 785-791.

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

Geczy, C.C., Stambaugh, R.F., Levin, D., 2003. Investing in socially responsible mutual funds. Working Paper Wharton School. Retrieved from: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=416380>. [12 May 2014].

Goldman Sachs – Goldman Sachs Global Investment Research (2007), Overview: Introducing GS SUSTAIN, Available from: <http://www.natcapsolutions.org/business-case/GoldmanSachsReport_v2007.pdf>. [13 May 2014].

Goldreyer, E.F. & J. D. Diltz (1999). The Performance of socially responsible mutual funds: incorporating socio-political information in portfolio selection. *Managerial Finance*, 25 (1), 23-36.

Gregory, A., Matatko, J., Luther, R., 1997. Ethical unit trust financial performance: small company effects and fund size effects. *Journal of Business Finance and Accounting*, 24 (5), 705–724.

Hendricks, D., Patel, J. & Zeckhauser, R. (1993). Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988, *The Journal of Finance,* 48 (1), 93-130.

Institutional Institute for Sustainable Development 2012, *Sustainable Development Timeline*. Retrieved from: < http://www.iisd.org/pdf/2012/sd_timeline_2012.pdf>. [2 May 2014].

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964, *Journal of Finance,* 23 (2), 389–416.

Jensen, M. C. (2002). Value Maximization, Stakeholder Theory, and the Corporate Objective Function. *Business Ethics Quarterly,* 12 (2), 235-256.

Lee, D. D., Humphrey, J.E., Benson, K.L. & Ahn, J.Y.K. (2010). Socially responsible investment fund performance: the impact of screening intensity, *Accounting & Finance,* 50 (2), 351-370.

Luther, R.G., Matatko, J., Corner, D., 1992. The investment performance of UK ethical unit trusts. Accounting, *Auditing and Accountability Journal*, 5 (4), 57–70.

Ma, C., Tchen, T., Smith, T., & MacNamara, A. (2011). The 'risky' risk-free rate: does the downgrade of US sovereign debt change commonly-used valuation approaches? *Financier Worldwide Limited*. Retrieved from: <http://www.hl.com/email/pdf/FW_Sep2011.pdf>. [17 April 2014].

Malkiel, B.G. & Xu, Y. (1997). Risk and return revisited, *Journal of Portfolio Management*, 23 (3), 9–14.

Mallin, C.A., Saadouni, B., Briston, R.J., 1995. The financial performance of ethical investment funds. *Journal of Business Finance and Accounting*, 22 (4), 483–496.

Markowitz, H.M. (1952), Portfolio Selection, *The Journal of Finance* 7 (1), 77–91.

Miles, S. (2012). Stakeholder: Essentially Contested or Just Confused? *Journal of Business Ethics*. 108 (3), 285-298.

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

Riedl, A. & Smeets, P. (2014). Social Preferences and Portfolio Choice. Retrieved from: <

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2318987>. [18 April 2014].

Schröder, Dr M. (2004). The performance of socially responsible investments: Investment funds and indices, *Financial Markets and Portfolio Management,* 18 (2), 122-142.

Sharpe, W.F. (1966). Mutual fund performance, *The Journal of Business,* 39 (1), 119-138.

Statman, M. (2000). Socially Responsible Mutual Funds, *Association for Investment Management and Research*. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.197.6247&rep=rep1&type=pdf>. [17 April 2014].

Swesif 2014, *Hållbarhetsprofilen*. Retrieved from: <http://www.hallbarhetsprofilen.se>. [8 April 2014].

The Generation Foundation (2013). *Sustainability and Investing*. Retrieved from: <http://www.generationim.com/sustainability/investing.html>. [18 April 2014].