UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Ethical Fund Performance

-A matched pair analysis of the Swedish fund market

Bachelor of Science (B.Sc.) in Finance

University of Gothenburg – School of Business, Economics and Law
Institution: Centre for Finance
Supervisor: Mari Paananen
Bachelor's Thesis Spring 2019
15 Credits

Jesper Lundgren and Robin Olin

Gothenburg, Sweden

Spring 2019

Abstract

This thesis investigates the effect of ethics on the performance of Swedish funds over the years 2009-2018. Through the use of environmental, social, and governance (ESG) score, this study distinguishes ethical funds from the less ethical funds. These funds are then compared and analyzed further with the help of traditional risk-adjusted performance measurements. For the final step of the thesis, these measurements, together with additional explanatory variables, were used to examine the ESG score effect on fund performance through a panel data regression. The findings show that fund performance is dependent on ESG score at the 5% significant level. However, the results of the study also suggest that there is a tendency of the less ethical funds outperforming their ethical counterpart. With regards to this and the fact that other findings of this study had non-significant numbers, no conclusions can be drawn about one group outperforming the other.

JEL Classifications: G11, G12, Q56
Keywords: ESG, Mutual funds, Swedish Fund Market

Acknowledgment

We want to thank our supervisor Mari Paananen, for her constructive comments and guidance throughout the thesis writing. We also would like to express a special thanks to Aineas Mallios for his further assistance when working with STATA.

Abbreviation

Capital Asset Pricing Model	-	CAPM
Environmental Social Governance	-	ESG
Net Asset Value	-	NAV
Ordinary Least Square	-	OLS
SIX Return Index	-	SIX RX
Social Responsible Investing	-	SRI
Stockholm Interbank Offered Rate	-	STIBOR
Sustainable Development Goals	-	SDG
Variance Inflation Factor	-	VIF

Table of Content

1. Introduction

In recent years ethical considerations have become an essential aspect for managers when making their investment decisions. So-called ethical investing has received increased influence together with the rise of technology and increased awareness of the environmental issues in the last couple of decades (El-Hagger, 2007). Another concept broadly discussed when investing ethical is the importance of sustainable investing (Hale, 2017). As of 2015, the United Nations implemented 17 goals for sustainable development. These goals, together with the Paris Agreement of 2016, hopes to stress the importance of sustainable actions taken all over the world, which also enlightens the importance of sustainable investing (UN.org, 2019).

When evaluating ethical investing further one could say that it is an investment strategy where personal values of social, moral, and religion are taken into consideration when creating a portfolio (Wealthsimple.com, 2019). There is no clear definition for the concept of ethics, but there are at least several ethical ratings acceptable on a larger scale used for measuring levels of ethics. Commonly used ratings that measure levels of ethics in businesses include social responsible investing (SRI) and environmental, social, governance (ESG) (ibid).

Prior research by Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005) exploits the associations between the impact of ethics on fund performance in the United Kingdom. Unlike these two studies, this thesis examines Swedish open-ended mutual funds, as well as comparing them against a benchmark. Sweden together with the other Nordic countries tops the 2018 rankings for the global Sustainable Development Goals (SDG) index which measures a countries performance in terms of fulfilling the 17 goals for sustainable development (Sachs, 2018). This study might get a different result than that of Mallin, Saadouni, and Briston (1995) who examined the U.K. market since U.K. were in 2018 rated in place 14 of the SDG index (ibid).

1.1 Purpose of the study

The purpose of this thesis is to evaluate the performance of Swedish open-ended mutual funds with the highest ESG score and examine whether they have a higher risk-adjusted performance relative to their matched counterparts. This thesis is similar to that of Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005). Furthermore, this study aims to extend Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005) by analyzing if funds ESG score is an explanatory variable for performance. The study will try to answer the

following research question, which is tied to the four hypotheses presented in section two of the thesis.

- Does the ESG score have any effect on the financial performance of Swedish mutual funds?

The study investigates the Swedish fund market post-financial crisis 2007/2008 and ten years onwards until the end of 2018. It aims to create value for an investor of how the ESG score can effect the returns of their portfolio.

1.2 Thesis structure

The rest of the thesis is structured as follows: Section two introduces a theoretical framework consisting of previous research and hypotheses that will be used throughout the thesis. In section three, the method of the thesis is presented, followed by section four presenting the data along with descriptive statistics. Section five consists of the research results and robustness test, while section six presents a discussion of possible limitations of the study as well as suggestions for future research topics. The thesis ends with a conclusion.

2. Theoretical Framework

This section examines our research question more closely with the help of theoretical aspects and previous research within the field of finance. In the hypothesis development, the four hypotheses for the thesis will be presented as well as several theories that were used to test the hypotheses further.

2.1 Previous research

The financial performance of funds has been the main subject in several previous studies. More specifically examining ethical and less ethical fund performance has been done in several different ways, but there are some similarities in how they proceeded. Our study examines several of the most frequently used concepts for examining fund performance; however, it will also contain an extension of these concepts.

Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005) are two of the most prominent studies on evaluating ethical and non-ethical fund performance. We follow Mallin, Saadouni, and Briston (1995) approach for the matching and comparing of the ethical and less ethical funds. We use the same matching criteria, age and size, for our sampled funds. One of the main differences between our study and prior studies is our proxy for ethical funds. Mallin, Saadouni, and Briston (1995) define ethical funds as to whether they fulfill negative criteria or positive criteria. Funds satisfying the negative criteria are those funds having policies not to invest in specific industries such as alcohol, tobacco, gambling, and so on (ibid). Differently, the positive criteria are those funds that invest in environmentally friendly companies (ibid). We characterize our sampled funds based on the ESG score of the funds. The ESG score helps an investor to consider the sustainable intentions of the chosen fund (Hale, 2017). The score is based upon how well the funds are addressing ESG issues such as those determining preparedness, disclosure, and performance of the funds, hence making it a reliable measurement for ethics in businesses (ibid).

Mallin, Saadouni, and Briston (1995) was conducted solely on the U.K. market, while Kreander et al., (2005) used funds from four European countries. Neither of the two studies found any evidence that ethical funds would outperform the market. On the contrary, Mallin, Saadouni, and Briston (1995), found evidence that the financial performance of the ethical funds tends to underperform both the non-ethical funds and the market. Their study also unveils that on a risk-adjusted basis, both groups tend to underperform the market (ibid). Kreander et al., (2005) found evidence suggesting that there was no difference between ethical and non-

ethical funds according to the performance measures used. Kreander et al., (2005) also extended Mallin, Saadouni, and Briston (1995), as they found evidence that neither the ethical nor the non-ethical funds had any ability to time the market. Sharpe (1975) express market timing as the strategy and ability to move in and out from financial markets and to switch between assets based on predictive methods.

2.2 Hypotheses development

The financial performance of funds differs because of several reasons. Hence, for a reliable comparison between ethical and less ethical funds, a matched sample analysis is formed. By using the funds size and the formation date of the funds as matching criteria, Mallin, Saadouni, and Briston (1995) state that it "[...] should help to eliminate the effect of specific characteristics which may be endemic in ethical investment funds' portfolios" (p. 484). Matched sample analysis eliminates characteristics such as the small company effect since ethical investments may more commonly occur in smaller companies (Mallin, Saadouni, and Briston, 1995). Our study also takes different types of investments made by the funds into consideration for the matching process. Since some of the most appropriate matched funds contain relatively high ESG score, hereafter the further analysis will be to examine ethical fund performance and less ethical fund performance. With the matching criteria in place, summarized in Appendix 1a, the geometric return of the top 30 ESG scored funds, and the 30 less ethical funds, are now comparable to the benchmark, enabling the study to test for the first null hypothesis stated below.

$H_{0,1}$: There is no difference in financial return between the two groups, and the benchmark.

The test for the first hypothesis will not tell us much about the actual performance of our sampled funds. Running more tests helps the study to get a significant analysis of differences in Swedish funds' performance. Therefore, the funds are analyzed further with the usage of different measurement techniques, most commonly used in financial research. Traditional risk-adjusted performance measures such as Capital Asset Pricing Model (CAPM), Jensen's alpha, Sharpe ratio and Treynor ratio are added to control for differences in risk-taking, similar to what Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005) did in their research.

CAPM is a financial model that explains the relationship between systematic risk and expected return of assets (Sharpe, 1964). The idea with the model is that the return on a portfolio is a result of the risk-free rate together with the excess return on the market. CAPM states that portfolios will only compensate for levels of market risk (ibid). Other assumptions for the CAPM is that all investors have the same time horizon, they are price-takers, and they are all rational and risk-averse (ibid). Neither are there any transaction costs or trade restraints, meaning that the investors can borrow an unlimited amount under the risk-free rate (ibid). The goal of the model is to examine if the asset is reasonably priced and whether it receives the appropriate excess return to compensate for risk and the time value of money (ibid). Sharpe (1964) claims that assets having a high correlation with the market will receive a higher expected return while less correlated assets will receive a lower expected return. The CAPM-model is as follows:

$$E[R_i] = R_i = R_f + \beta_j(E[R_{mkt}] - R_f) + \epsilon \quad (1)$$

$E[R_i] =$ expected return for the portfolio

$\beta_j = \frac{cov(\tilde{R}_j, \tilde{R}_M)}{\sigma^2 \tilde{R}_M} =$ Systematic risk

$E[R_{mkt}]$ = expected return on the market

$R_f =$ Risk free rate of interest

$\beta(E[R_{mkt}] - R_f) =$ Risk premium for security i

$\varepsilon =$ Error term

Jensen's alpha is a risk-adjusted performance measure that represents the average return of a portfolio, or an investment, compared to the expected return of that portfolio (Jensen, 1968). Jensen's alpha measures and evaluates portfolio managers predictive ability (ibid). That is, her ability to earn returns through successful prediction of security prices, which are higher than expected, given the level of riskiness of the portfolio (ibid).

$$\alpha = R_i - [R_f + \beta_j\big(E[R_{mkt}] - R_f\big)] \quad (2)$$

The Sharpe ratio is an evaluation measure of portfolios past performance (Sharpe, 1966). The ratio is designed to measure the expected excess return per unit of risk (ibid). Sharpe (1964) claims that when allocating funds among several funds, the Sharpe Ratio will provide useful guidance and that "[...] it makes sense to favor the one with the greatest predicted Sharpe

ratio [...]" (p.16). According to the model, the risk level of the funds will be proportional to the predicted Sharpe ratios for the selected portfolio returns (Sharpe,1994). Therefore, the Sharpe ratio is a good measure to use when studying the risk level of the matched funds.

$$SHARPE\ RATIO\ = \frac{\tilde{r}_j - \tilde{r}_f}{\sigma_j} \quad (3)$$

$\tilde{r}_j$ = Return on the market

$\tilde{r}_f$ = Risk free rate of interest

σ_j = Standard deviation of the fund

Treynor ratio is a similar measurement technique to that of the Sharpe ratio but instead of basing its calculations on the standard deviation of the portfolio; it uses the beta of the fund. It is a risk-adjusted performance measure of return based on systematic risk developed by Treynor (1965). The ratio indicates how much excess return an investment has earned in respect to the level of risk the investment contained (ibid).

$$TREYNOR\ RATIO\ = \frac{\tilde{r}_j - \tilde{r}_f}{\beta_j} \quad (4)$$

$\tilde{r}_j$ = Return on the market

$\tilde{r}_f$ = Risk free rate of interest

β_j = Beta of the fund

With the risk-adjusted performance measure applied on all of our sampled funds, the study could now test for the second null hypothesis below.

$H_{0,2}$: There is no difference in risk-adjusted financial performance between ethical funds, less ethical funds, and the benchmark.

Thus far, the thesis has only compared the two groups of funds without any further comparisons between each matched fund. Therefore, the study tests for the third null hypothesis stated below.

$H_{0,3}$: There is no difference in financial performance between each matched pair of the ethical and less ethical fund.

The third hypothesis is an extended version of previously made studies by Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005). Prior sections of this study have examined funds financial returns with and without adjusting for risk and whether the funds alphas are statistically significant or not. This part the study tests whether there is a significant difference in financial performance individually between the funds with the highest ESG score and their matched counterpart. By conducting a two-sided t-test, the study tests if there is a significant difference between the means of each matched fund (Student, 1908). For example, if fund 1A differs significantly from 1B, shown in Appendix 1a. The test produces a t-value as its output. A low t-value indicates that there is no difference between the two means of the funds and vice versa for a high t-value.

For the second hypothesis, several risk-adjusted performance measurements were introduced to be able to control the level of risk taken by each fund. The third hypothesis does not include the before-mentioned risk-adjusted measurements used for testing hypothesis two; however, one could still argue that the test for hypothesis three adjusts for the appropriate risk taken by investors. In models based upon rational fund manager behaviors, the investors of the funds have already taken the level of risk associated with investing in the fund into consideration.

The fourth and final hypothesis of the thesis tests whether or not ESG score is an explanatory variable for determining fund returns. Since the ESG score distinguishes the level of ethics for each fund, it is interesting to test whether or not fund returns are dependent on the ESG score. Following is the fourth null hypothesis of the thesis.

$H_{0,4}$: The financial return of funds is not dependent on the ESG score.

In order to test for the fourth hypothesis, we introduce a regression analysis where the ESG score is an explanatory variable for fund returns.

$$RETURN = \beta_0 + \beta_1 ESGhigh + \beta_2 x2 + \cdots + \beta_k x_k + U \quad (5)$$

3. Method

This section describes the method of the constructed tests for each of the four hypotheses. The method presents the research design and how the study tests the hypotheses developed in the theoretical framework.

The ESG scored used throughout the study was retrieved from the Morningstar website. The ESG score is composed by the research firm Sustainalytics, who is considered to be one of the leading research firms in the industry (Morningstar.com, 2016). Both the data containing the size of the funds as well as their appropriate ESG score is as of March 31st, 2019. As the funds are primarily matched by the date of creation, the study is based on the size of the funds today. Furthermore, the ESG score is also assumed to be constant over the studied time-period due to absence of ESG score from the earlier years. However, the ESG score is based upon historical holdings of the funds which makes the score more reliable (Morningstar.com, 2018).

Once the entire sample of funds were matched, the net asset value (NAV) for each fund and the benchmark were downloaded from the Bloomberg terminal. The following formula displays how the monthly averages of the NAV were calculated to be able to test for the first hypothesis.

$$Monthly\ return\ = \frac{NAV_T - NAV_{T-1}}{NAV_{T-1}} \tag{6}$$

NAV_T = Net asset value this month

NAV_{T-1} = Net asset value last month

With downloaded NAV, the performance of the funds and the benchmark was measured by calculating the historical financial return of the funds. A ten-year average of the STIBOR 3 month was used as the risk-free rate. With the usage of STATA, the calculations of each funds historical financial return were calculated and compared against the chosen benchmark.

The second hypothesis is an extended test of the first hypothesis. Here the performance of the funds is analyzed further once accounted for the level of risk taken by the funds. The second hypothesis uses the annual means of the funds to calculate the risk-adjusted performance measurements properly.

The following section describes the conducted calculations for the Sharpe ratio, Treynor ratio, and Jensen's alpha. The first risk-adjusted measurement used in the study was the Sharpe ratio, equation 3 seen in section 2.2. With the return of the funds calculated in hypothesis one,

the risk-free rate was deducted to retrieve each funds risk premium. The risk premium was then divided by the standard deviation for each fund which was withheld from STATA. The second risk-adjusted measurement used in this study is the Treynor ratio, equation 4 seen in section 2.2. In order to calculate the Treynor ratio of each fund, the risk premium of the funds was divided with the beta of that specific fund. The beta values used for the Treynor ratio were retrieved from the CAPM, equation 1 stated in section 2.2. To enable the study to calculate values from CAPM, the risk-free rate, as well as the benchmark, had to be retrieved from the Bloomberg terminal. With the help of the risk-free rate and the benchmark, we calculate the beta of the funds and the Jensen's alpha, equation 2 seen in section 2.2.

The third hypothesis uses the returns of the funds from the previously stated hypotheses. Unlike previous tests, this section tests each matched pair individually against each other. The matching process, as described before, was done by inception date, fund size, and with respect to investments made by the funds. In order to examine whether the two funds historical financial return differ at a statistical or marginal significance level, a t-test was formed in STATA.

When testing for the fourth and final hypothesis of the study, the data had to be rearranged into a panel data format since the thesis is based upon several funds and their returns across time. The primary goal with the regression model created for the fourth hypothesis is to evaluate the ESG scores actual impact on the return of the funds. In order to get a better comparison, the panel data was clustered by company. Due to many different levels of ESG scores for our sampled funds, the ESG score is treated as a dummy variable in the regression so that the treatment group consisting of the top 30 ethical funds could be separated from the control group consisting of 30 less ethical funds. The risk-adjusted performance measurements used for testing hypothesis two were also added as regressors to properly evaluate the ESG score when the level of risk is taken into account. Other variables that the model controls for is the risk premium, the natural logarithm of asset size, as well as checking for industry effects. By using dummy variables, the funds were separated into five industries, in which the funds to be able to test for the different investments separately. Manufacturing, healthcare, technology, energy, and financial services are the five industries in which the funds had their largest holdings. We also added interaction terms for the combination of ESG high and industry to evaluate the ESG score effect on return in different industries. Since it is crucial to detrend variables that may be affected by past performances and ensure that the model no longer experienced high persistency, a test for correlation was made for time lag of the dependent variable (Kendall, 1956). After testing for correlation, seen in Appendix 5b, the depended time lag is of no use.

To create a reliable regression model, the six Ordinary Least Square (OLS) assumption expressed below had to be fulfilled. The assumptions tell us that the regression has to be linear, and each regressor needs to be exogenous (Wooldridge, 2014). Furthermore, no variable is allowed to be linearly dependent on the other since that would lead to exact multicollinearity and consequently break the full rank assumption (ibid). The full rank assumption is tested by running a Variance Inflation Factor (VIF) test in STATA to ensure that the model is free of multicollinearity. After running the VIF test, seen in Appendix 5c, the interaction term manufacturing and ESG high, as well as the risk-adjusted performance measurement Sharpe ratio and Treynor ratio had multicollinearity and were excluded from the regression. When testing the Sharpe ratio and Treynor ratio individually without including the Jensen's alpha, they still experienced multicollinearity. Therefore, no extra regressions were tested with those two measures included. The VIF test with all the variables used can be seen in Appendix 5d. The two interaction terms that included the industries energy and technology were excluded from the regression because neither of the funds in the treatment group had their largest holding in those industries.

Other critical OLS assumptions are that the variables are collected through random sampling; the sample has homoscedastic error terms and that the sample is normally distributed (ibid). To ensure that there was no heteroscedasticity in the model, robust standard errors were used. Since the data is from time-series, there might be trends causing autocorrelation, which also needs to be taken into account (Kendall, 1953). When considering all the OLS assumptions, the thesis regression is as follows:

$$RETURN = \beta_0 + \beta_1 ESGhigh + \beta_2 RP + \beta_3 RADJ + \beta_4 Size + \beta_{5-9} IND + \beta_{10-11} IT + U \quad (7)$$

$ESGhigh$ = Being part of the treatment group
RP = Risk premium
$RADJ$ = Jensen's alpha
IND = Industry
IT = Interaction term
β = Coefficient of the variable

4. Data

This section presents the collection of the data used in the thesis. It contains descriptions of the data and how it was used and processed to analyze the historical performance of the funds as well as their level of risk. Explanatory variables such as the benchmark and risk-free rate as well as other essential definitions of concepts used in this thesis will be explained more thoroughly in this section.

4.1 Sample selection

When conducting the study, there was a total of 277 open-ended mutual funds with Swedish domicile existing during the period 2009-2018. These funds were narrowed down to 30 funds with the highest ESG score. Later the 30 funds were matched to create a control group consisting of 30 less ethical funds. As previously stated, the matched sample approach was conducted similar to the ones made by Mallin, Saadouni, and Briston (1995), Kreander et al., (2005) and Gregory, Matatko, and Luther (1997) (Sample Shown in Appendix 1b) as this study controls for fund size, age, and investment holdings. The size of the funds was retrieved from the Bloomberg terminal. Due to the limited population of Swedish domiciled open-ended mutual funds, only eight funds were matched with regards to their investment holdings. To be able to investigate the performance of mutual funds, monthly NAV, as well as the size of the funds, were obtained from the Bloomberg terminal. The data fulfilling the matching criteria were all collected from this terminal. The Morningstar website was another database used to be able to get the correct ESG score for our sampled funds.

We use the ESG score as the measurement for distinguishing ethical funds from less ethical funds. In more recent years, the ESG score has become the most common and reliable rating for the level of ethics in funds (Hale, 2017). The ESG score is used by investors to be able to consider environmental, social, and governance issues when investing (Morningstar.com, 2016). As of 2018, in order for a fund portfolio to receive an ESG score, at least 67% of the fund holdings must have an ESG score (Morningstar.com, 2018). A fund portfolio ESG score is an asset-weighted average of normalized company-level ESG scores (ibid). To make the ESG score comparable across peer groups, Morningstar (2018) calculates a z-score which creates a normalized ESG score on a 0-100 scale. On the next page is Morningstar (2018) interpretation of the ESG score.

70 + = Company scores at least two standard deviations above average in its peer group.

60 = Company scores one standard deviation above average in its peer group.

50 = Company scores at peer group average

40 = Company scores one standard deviation below average in its peer group

30 - = Company scores at least two standard deviations below average in its peer group.

Table 1 consist of several examples of ESG issues relevant for determining the ESG score.

Table 1 Examples of ESG issues

Environmental issues	**Social issues**	**Governance issues**
Climate change and carbon emissions	Customer satisfaction	Board composition
Air and Water pollution	Data protection and privacy	Audit committee structure
Biodiversity	Gender and diversity	Bribery and corruption
Deforestation	Employee engagement	Executive compensation
Energy efficiency	Community relations	Lobbying
Waste management	Human rights	Political contribution
Water scarcity	Labor standards	Whistleblowers schemes

Source: Environmental, Social, and Governance Issues in Investing: A Guide for Investment Professionals, CFA Institute

This study does not only compare performance and returns for ethical and less ethical holdings, it also contains valid information and comparisons to a benchmark. We use the SIX Return Index (SIX RX) as a benchmark. SIX RX measures the performance of all companies listed on the Stockholm Stock Exchange with their dividend included. Therefore, SIX RX becomes a more precise benchmark compared to OMXS30, which only contains the top 30 most traded stocks on the Stockholm Stock Exchange (Matilainen, Petersson and Eriksson, 2012). Since the study examines diversified portfolios, it is relevant to choose a benchmark like SIX RX that shows the stock returns of the companies when the dividend is included. With the SIX Return Index, the thesis gets a benchmark that is well diversified and consists of a large sample of companies which is beneficiary for the matching process.

The risk-free rate used throughout the thesis is the Stockholm Interbank 3 Month Offered Rate, or as it is also called, the STIBOR. The STIBOR shows the interest rate at which

seven large Swedish banks are willing to lend money to each other for three months without security (Riksbanken.se, 2018). The STIBOR used for the thesis is the ten-year average of the investigated period.

4.2 Descriptive statistics

Table 2 presents descriptive statistics of the treatment group and the control group. It is a summary of the data collected from the entire sample period from January 2009 until December 2018. The table is divided into three rows, presenting the monthly average performance of the sampled funds along with a summary of the difference between them. The first column shows the geometric means of the funds with the highest ESG score together with the funds matched counterpart. The geometric mean is used for the calculation of the average return of the funds, since it considers the cumulative return of the funds. As table 2 shows, the treatment group has a higher geometric mean return, suggesting that they generate more substantial returns than that of the control group. As a supplement to the mean values the 25th percentile, the median, and the 75th percentile is presented to show whether there is any skewness of the sample in any particular direction. The minimum and maximum values show if any extreme outliers affect the mean, and as noticed the Control group has a larger spread between minimum and maximum values. The seventh column presents the standard deviation of the groups. As seen, the treatment group has a slightly higher standard deviation than the control group suggesting more significant risks associated with investing in the top 30 ethical funds.

Table 2 Descriptive Statistics

	Mean	*25th*	*Median*	*75th*	*Min*	*Max*	*St. Dev*	*No. Obs*
Treatment Group	*0.0094*	*-0.0130*	*0.0110*	*0.0309*	*-0.1508*	*0.2752*	*0.0416*	*3,600*
Control Group	*0.0088*	*-0.0134*	*0.0088*	*0.0305*	*-0.1618*	*0.3198*	*0.0414*	*3,600*
Difference	*0.0006*	*0.0004*	*0.0022*	*0.0004*	*0.0110*	*-0.0446*	*0.0002*	

Notes:
(i) 25^{th} and 75^{th} denotes the percentile

5. Empirical Results

In this section the result of the study is presented. The chapter starts with the results from each hypothesis, followed by a test of robustness. The section contains detailed descriptions of how models were applied to derive the results of the thesis.

When testing the first hypothesis, the historical financial returns of the treatment group and the control group were compared to the benchmark. Appendix 2a shows the monthly average geometric return of each fund as well as the benchmark average. As table 3 shows, the treatment group underperformed the benchmark in 28 instances while the control group underperformed in 26. In two respectively four cases, the sampled funds outperformed the market. This result is in line with the efficient market hypothesis created by Fama (1970) and his discussion about the difficulty of beating the market. It is also supported by more recent studies of European funds suggesting that more than 86% of the mutual funds and approximately 79% of Swedish equities underperformed against the market during the last ten-year period (Cairns, 2019) The result presented in Appendix 2a shows that both groups of funds experience larger standard deviations than the benchmark, indicating considerable dispersions for the treatment group and control group than that of the benchmark. A high standard deviation tells us that the fund has high volatility, which is associated with greater risk.

Table 3 Financial Return Comparison

	Treatment group		**Control group**	
	Underperform	Outperform	*Underperform*	Outperform
Financial Return 2009-2018	28	2	26	4

In order to extend the comparison between the two groups and answer the second hypothesis, an analysis of the risk-adjusted performance was arranged. A summary of the results from the Sharpe ratio, the Treynor ratio, and Jensen's alpha is shown in table 4. The control group has a higher Sharpe ratio in 17 of 30 cases with a mean of 0.71182 compared to 0.70496 for the treatment group. This data indicates that the control group experience higher returns for the given level of risk compared to the treatment group. However, the difference between the means of the two groups, 0.00686, is not significant at the 5% level as seen in Appendix 4a. Neither is the median of the Sharpe ratio significant at either the 5%, or 10% level. The result from the Treynor ratio shows more substantial variations between the two groups of funds, seen in Appendix 4b. The average Treynor ratio of the treatment group is

0.11293, while the average of the control group is 0.14550. As explained in the theory section, there is a close resemblance between the Sharpe ratio and the Treynor ratio with the difference that the Treynor ratio uses the beta instead of the standard deviation in the denominator of the equation. A high Treynor ratio indicates that given the level of correlation to the market, the return of the investment is more significant than if the ratio were to be small. The mean difference of the Treynor ratio between the two groups is 0.03257, and it is significant at the 5% level along with it being greater in 21 out of 30 cases for the control group. Worth mentioning is that when looking at the median of the two groups, the difference is smaller and not significant at either the 5%, or 10% level. This implies that some of the funds in the control group experience higher Treynor Ratios and hence, affects the mean variable. Something that is also visible in Appendix 2b, where the individual statistics of the funds are shown.

Table 4 Risk-Adjusted Comparison

	Sharpe			**Treynor**			**Jensen's alpha**		
		Mean	*Median*		*Mean*	*Median*		*Mean*	*Median*
Treatment Group	**13**	*0.705*	*0.736*	**9**	*0.113*	*0.114*	**11**	*-0.002*	*-0.001*
Control Group	**17**	*0.712*	*0.735*	**21**	*0.146**	*0.120*	**19**	*-0.001*	*-0.001*

Notes:
(i) * denotes significance at the 0.05 level or better

When analyzing the third and final risk-adjusted performance measure, Jensen's alpha, nearly all open-ended mutual funds underperform the market. The average alpha of our total sample of 60 funds is -0.0015104. Neither of the funds in this sample had significantly positive alphas. On the contrary, there are nine funds in the treatment group and four funds in the control group that had negative alphas at the 5% significance level. Once again, implying the difficulty of beating the market (Fama, 1970). Four additional funds from both groups also received negative alphas, but they were only marginally significant at the 10% level. The average alpha for the funds in the treatment group was -0.001813 and the control group had an average alpha of -0.001208 but with no statistical difference between the means of the two groups at the 5% significance level. Neither is the median of the Jensen's alpha significant at 5% level. Some 11 funds in the treatment group had higher Jensen's alpha values than the funds in the control group.

Table 5 Jensen's Alpha Significance

	Positive	Significance level		**Negative**	Significance level	
		5%	*10%*		*5%*	*10%*
Treatment Group	**2**	*0*	*0*	**28**	*9*	*4*
Control Group	**4**	*0*	*0*	**26**	*4*	*4*

The third hypothesis further expands prior research by Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005) through individually comparing the matched funds. The results of the individual comparison are presented in Appendix 3a, where all T-values from the compared funds are displayed. There are only four instances where the T-value of the compared funds are significant at the 5% level. For the funds to differ significantly at the 5% level, the T-value has to be higher than the absolute value 1.96, which is equal two standard deviations from the mean, assuming a normal distribution. Four comparisons contained such T-values indicating significant differences in means between the compared funds. Of these four, two were from the treatment group, while two of them were from the control group.

As the final part of the thesis, a panel data regression was conducted to test the fourth hypothesis of whether the financial return is affected by the ESG score of the funds. The output from STATA, seen in Table 6, shows that the risk premium was significant at the 5% level while the size of the funds was not. Not shown in table 6 is the result from the industry dummy variables, that were positively significant at the 5% level for both manufacturing and the interaction term, healthcare and ESG high. The variable energy was negative and significant at the 5% level. Table 6 also shows that the variable ESG high is positive and significant at the 5% level. This result indicates that the ESG score is affecting the financial performance of the funds in a positive way and therefore we can reject the fourth null hypothesis.

Table 6 *Regression Analysis*

	Return
ESG high	4.4 e-05*
	(2.02)
Risk premium	0.938*
	(306.05)
Jensen's alpha	0.057*
	(5.36)
Size	9.38e-07
	(0.14)
Constant	1.007*
	(3.7e+04)
Industry Dummies	Yes
Observations	7,200
Overall R-squared	0.9409
Number of Funds	60
Period	08-18

Notes:
(i) t-values are in parentheses.
(ii) * denotes significance at the 0.05 level or better.

5.1 Robustness Test

The following section presents a robustness test. For the robustness test, a shorter period for the sample was investigated to exclude possible economic turbulence caused by the financial crisis of 2008. The Swedish economy was not out of the recession until 2015, following the devastating financial crisis in 2008 (Konj.se, 2015). Consequently, it would be interesting to exclude those years where the Swedish economy was still in a recession and test our hypotheses once again for the time-period, 2015-2018.

When applying the test of robustness for a shorter time-period on the first hypothesis, we found that both the ethical funds and the less ethical funds outperformed the market on more occasions, which can be due to many reasons. One reason could be the increased health of the Swedish economy for the tested time-interval, which may have caused the stronger returns of the two groups. Another reason may exist due to "the random walk theory" stressing the difficulties with performing better than the market over a more extended period of time due to

the random walk of stock prices (Malkiel, 1999). The results of the shorter, as well as the longer period, are shown in table 7 below.

Table 7 Robustness Test Financial Return Comparison

	Treatment group		**Control group**	
	Underperform	*Outperform*	*Underperform*	*Outperform*
FinancialReturn 2015-2018	21	9	16	14
FinancialReturn 2009-2018	*28*	*2*	*26*	*4*

For the second hypothesis, some differences emerged when performing the test of robustness as well. As seen in table 8, the average alphas of the treatment group were still negative, but the funds in the control group now had positive alphas, as well as having a statistically significant difference in means between the two groups at the 5% significance level. This result indicates that the fund managers with responsibility of the less ethical funds performed better in recent years than for the whole period. Once again, the health of the Swedish economy may have had an impact on the findings of the robustness test.

Table 8 Robustness Test Jensen's alpha

			Positive	Significance level		**Negative**	Significance level	
	Mean	*Median*		*5%*	*10%*		*5%*	*10%*
Treatment Group	*-0.001*	*-0.001*	**12**	*0*	*0*	**18**	*3*	*1*
Control Group	*0.002*	*0.002*	**24**	*3*	*2*	**6**	*0*	*0*

As part of the robustness test for the third hypothesis, no further conclusions arose from the already existing empirical result.

The robustness test of the fourth hypothesis showed a few similarities yet a few differences from the empirical results. The output in table 9 showed that for a shorter period, the variable ESG high was still positive, but this time not significant at either the 5%, or 10% level. Both the Jensen's alpha and the combination of ESG high and healthcare was, on the contrary to previous results, negative at the 5% significance level. Furthermore, the variables energy, financial services, and healthcare were now significant at the 5% level. The results of the robustness test show the difficulties of finding a reliable variable that through time has a constant effect on financial return. Some uncertainties arose from the findings of the robustness test. As seen in table 9 the variable, ESG high, is no longer significant indicating that for the

shorter time-period ESG is no longer an explanatory variable for fund performance. Even though environmental issues as well as overall sustainable knowledge have had increased importance in recent years, none of this can be shown in our findings of the robustness test. Therefore, when analyzing our main findings together with our findings in the robustness test, we speculate that even though people may be more environmental aware the last four years, this may not have any impact on how ESG scores affect funds financial performance. Table 9 below presents the result of the robustness test for the fourth hypothesis.

Table 9 Robustness Test Regression analysis

	Return
ESG high	1.80 e-07
	(0.10)
Risk premium	1.004*
	(1573.31)
Jensen's alpha	- 0.003*
	(-3.66)
Size	1.23e-07
	(0.20)
Constant	0.996*
	(2.0e+05)
Industry Dummies	Yes
Observations	2,800
Overall R-squared	0.9979
Number of Funds	60
Period	15-18

Notes:
(i) t-values are in parentheses.
(ii) * denotes significance at the 0.05 level or better.

6. Discussion

Our discussion section contains a critical examination of the method execution as well as an explanation of the main findings of the study. The models used for testing the expressed hypotheses, as well as the research findings, and test of robustness will be treated and analyzed further, together with used literature. Moreover, the possible shortcoming of the study will be discussed as well as suggestions for future research topics.

When analyzing the result, some characteristics of differences in financial performance between ethical and less ethical funds were discovered. However, due to many numbers being non-significant at either the 5%, or 10% level, the second null hypothesis, stating "there is no difference in risk-adjusted performance", and the third null hypothesis, stating "there is no difference in financial performance between each matched pair", could not be rejected. When looking at our empirical results, we can reject the first null hypothesis, stating "there is no difference in financial return compared to the benchmark", and the fourth null hypothesis, stating "the financial return is not dependent on ESG score".

The results of the test for the second and third hypothesis shows that there is an indication of the less ethical funds outperforming the ethical funds for both of the risk-adjusted measurements, Sharpe ratio and Treynor ratio. However, it is only for the Treynor ratio that the comparison between the means of the two groups is significant at the 5% level. These results indicate that, given the correlation to the market, there is a tendency of the less ethical funds experiencing better risk-adjusted returns then the ethical funds. When looking at the findings from hypothesis one and the Jensen's alpha findings in hypothesis two, these results are in line with the efficient market hypothesis, as both the ethical and less ethical funds underperform on average against the market. The only difference between the tests of the two hypotheses is that we can reject the first null hypothesis while we cannot reject the second null hypothesis due to insignificant numbers.

The result of the fourth hypothesis shows that the financial return of our sample of Swedish open-ended mutual funds is dependent on the ESG score. Hence, we can reject the fourth null hypothesis. When going back to the stated research question presented in the purpose of the thesis, the conducted study shows tendencies that the ESG score could effect the financial performance of Swedish open-ended mutual funds. The result of the fourth hypothesis implies that a high ESG score has a small, but yet positive, contribution to the financial return on the Swedish fund market. However, the study also suggests that there are other significant variables tested in the second hypothesis that implies less ethical funds outperforming ethical funds.

As we analyze the findings of all hypotheses in combination with the result of the robustness test, which showed significant difference in Jensen's alpha between the groups and that the variable "ESG high" no longer were significant, it is difficult to draw any further conclusions. Our results are different from that of Mallin, Saadouni, Briston (1995) since we cannot conclude that ethical funds outperform the non-ethical funds on a risk-adjusted basis. Instead, our results are more in line with Kreander et al., (2005) that suggest that there is no significant difference between ethical and non-ethical funds.

This thesis has shown the difficulty of proving how potential differences between levels of ethics impact the overall performance of the funds. That we could reject the first hypothesis was expected and in line with financial theory that discusses the difficulties of beating the market (Fama, 1970). A valuable contribution to the research field is that we can reject the fourth null hypothesis, showing that the ESG score is an explanatory variable for fund performance on the Swedish fund market. The findings of our study create further discussion of why not always do ethical investing.

The result of our study relies on rather strong assumptions for the sample collection process. These assumptions are the most significant shortcomings of the study since they may have misleading consequences for the outcome of the results. The most critical assumption made throughout the study is that the ESG score and size of the funds are constant over time Since both the ESG score and the size of the funds were collected on March 31st, 2019, the results could be somewhat misleading. Even though the ESG score for the sample of funds is based upon historical holdings, our study still has to assume that the present top 30 ESG scored funds have had higher ESG scores than their matched less ethical counterparts for the entire period investigated.

One of the matching criteria for creating a treatment group and a control group were to have similar size of the funds in each group. The size of the funds is also assumed to have been constant over the investigated period. This assumption creates flaws of the given result since there is a possibility that some of the matched funds in this study were not as similar in size throughout the entire investigated period. Another shortcoming of the conducted study is that it relies primarily on two major studies evaluating ethics impact on fund performance. Both Mallin, Saadouni, and Briston (1995) and Kreander et al., (2005) may be outdated, and the course of actions of their studies might not be as applicable on evaluating the fund market today.

With the complete study at hand, some suggestions for future research would be to correct the shortcomings of this thesis and further extend the fourth hypothesis. This could be done by using a larger sample, and evaluate all mutual funds with Swedish domicile, as well as

not using a matched pair analysis. If possible, future research should use weekly data with continuous asset size and ESG scores to exclude the assumption of constant size and ESG score for the tested period. We believe that with a greater number of observations, not assuming constant size and ESG scores, the regression would more accurately capture the effect of ESG score on fund performance. Another future research suggestion is to use the extended Carhart's four-factor model, which is developed from the CAPM and Fama French three-factor model. With the added Carhart's four-factor model, the result of our study could become more precise since the model accounts for the momentum effect of assets. Carthart (1997) says that assets tend to continue on a given path as it rises or falls.

7. Conclusion

This thesis studied the performance of 60 Swedish funds using several different performance measurements. Just as Mallin, Saadouni, Briston (1995) and Kreander et al., (2005) this study used a matched pair analysis to evaluate the performance of the 30 funds with the highest ESG score and their less ethical counterpart. The results from the first hypothesis, suggests that funds underperform the benchmark in 54 out of 60 cases which is in line with previous research made by Mallin, Saadouni, Briston (1995) and Kreander et al., (2005). The other findings of this thesis suggest that there are no significant differences between the two groups, and we cannot reject either the second or third hypothesis. We could however, at the 5% level, reject the fourth hypothesis suggesting that ESG score is a positive explanatory variable for financial return on the Swedish fund market.

8. References

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

El-Hagger, S. (2007). *Sustainable Industrial Design and Waste Management*. United States: Elsevier Inc.

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

Gregory, A., Matatko, J. and Luther. R. (1997). Ethical Unit Trust Financial Performance: Small Company Effects and Fund Size Effects. *Journal of Business Finance & Accounting*, 24(5), pp. 705-725.

Hale, J. (2017). *The Morningstar Sustainability Rating.* [pdf] Morningstar. Available at: https://www.morningstar.com/lp/sustainability-rating [Accessed 15 May. 2019].

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

Kendall, M.G. (1953). The Analysis of Economic Time-Series—Part I: Prices. *Journal of the Royal Statistical Society*, 116(1), pp. 11-34.

Konj.se, (2015). *Konjunkturläget Juni 2015.* [pdf] Stockholm: Konjunkturinstitutet. Available at: https://www.konj.se/download/18.42684e214e71a39d0723181/1436517282135/Konjunkturlaget-juni-2015.pdf [Accessed 15 May. 2019].

Kreander, N., Gray, R.H., Power, D.M. and Sinclair, C.D. (2005). Evaluation the Performance of Ethical and Non-Ethical funds: A Matched Pair Analysis. *Journal of Business Finance & Accounting*, 32(7-8), pp. 1465-1493.

Malkiel, B.G. (1999). *A random walk down Wall Street: including a life-cycle guide to personal investing*. New York: Norton.

Mallin, C.A., Saadouni, B. and Briston, R.J. (1995). The Financial Performance of Ethical Investment Funds. *Journal of Business Finance & Accounting*, 22(4), pp. 483-496.

Matilainen, H., Petersson, M. and Eriksson, M. (2012). *SIX Return Index.* 1st ed. [pdf] Stockholm: SIX Financial Information. Available at: https://www.six-group.com/financial-information/dam/downloads/products_solutions/six-index/market/six-return-index.pdf [Accessed 4 Apr. 2019].

Morningstar.com, (2016). *The Morningstar Sustainable Investing Handbook.* [pdf] Available at: https://www.morningstar.com/content/dam/marketing/shared/Company/Trends/Sustainability/Detail/Documents/Morningstar-Sustainable-Investing-Handbook.pdf [Accessed 4 Apr. 2019].

Morningstar.com, (2018). *The Morningstar Sustainable Raiting.* [pdf] Chicago: Morningstar Available at: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/744156_Morningstar_Sustainability_Rating_for_Funds_Methodology.pdf [Accessed 15 May. 2019].

Riksbanken.se, (2018). *Svenska Marknadsräntor.* [online] Available at: https://www.riksbank.se/sv/statistik/sok-rantor--valutakurser/forklaring-till-serierna/svenska-marknadsrantor/ [Accessed 4 Apr. 2019].

Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., Fuller, G. (2018): SDG Index and Dashboards Report 2018. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).

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

Sharpe, W.F. (1966). Mutual Fund Performance. *The Journal of Business*, 39(1), Part 2: Supplement of Security Prices, pp. 119-138.

Sharpe, W.F. (1975). Likely Gains from Market Timing. *Financial Analysts' Journal*, 31(2), pp. 60-69.

Sharpe, W.F. (1994). Sharpe Ratio. *The Journal of Portfolio Management,* 21(1), pp. 49-58.

Student. (1908). The Probable Error of a Mean. *Biometrika,* 6(1), pp. 1-25.

Treynor, J.L. (1965). How to rate management of investment funds. *Harvard Business Review*, 43(1), pp. 63-75. 13p.

UN.org, (2019). The Sustainable Development Agenda. [online] Available at: https://www.un.org/sustainabledevelopment/development-agenda/ [Accessed 18 May. 2019].

Wealthsimple.com, (2019). *What is Ethical Investing & Why Do it.* [online] Available at: https://www.wealthsimple.com/en-us/learn/what-is-ethical-investing [Accessed 5 May. 2019].

Wooldridge, J.M. (2014). *Introductory Econometrics*. 5th ed. United States: Cengage Learning.

9. Appendix

Table 1a

Swedish top 30 ESG ranked open-ended mutual funds and the matched sample of Swedish less ethical open-ended funds

Treatment group			*Control group*		
Fund Ref.	Name of Mutual funds	*ESG*	*Fund Ref.*	Name of Mutual funds	*ESG*
(1A)	Öhman Etisk Index Europa	65.34	(1B)	Nordea Avtalspensionsfond Midi	57.54
(2A)	Öhman Etisk Index Sverige	64.03	(2B)	SEB Swedish Value Fund	57.82
(3A)	Länsförsäkringar Europa Aktiv	63.65	(3B)	Handelsbanken Latinamerikafond	52.11
(4A)	AMF Aktiefond Europa	63.64	(4B)	Swedbank Robur Europafond Mega	60.09
(5A)	Cicero Focus A	63.51	(5B)	Guide Aktiefond Global	56.30
(6A)	Aktiespararna Topp-Sverige	63.04	(6B)	SPP Aktieindexfond Japan	52.01
(7A)	SPP Aktieindexfond Europa	62.96	(7B)	Swedbank Robur Mixfond Pension	--
(8A)	Skandia Idéer för livet	62.80	(8B)	SEB Österuropafond	48.44
(9A)	Swedbank Robur Ethica Sverige MEGA	62.77	(9B)	Nordea Generationsfond 80-tal	--
(10A)	Swedbank Robur Ethica Sverige	62.76	(10B)	Lannebo Mixfond	59.02
(11A)	Swedbank Humanfond	62.75	(11B)	SEB Nordamerikafond	51.35
(12A)	Handelsbanken Europafond Index	62.68	(12B)	Swedbank Robur Medica	57.20
(13A)	Lannebo Sverige	62.55	(13B)	Öhman Etisk index USA	57.32
(14A)	Handelsbanken Nordenfond	62.55	(14B)	Nordea Allemansfond Alfa	60.15
(15A)	SPP Aktiefond Sverige	62.55	(15B)	Swedbank Robur Globalfond Mega	51.97
(16A)	Skandia Europa Exponering	62.54	(16B)	SEB Sverige Stiftelsefond	58.73
(17A)	Guide Aktiefond Sverige	62.32	(17B)	Spiltan Aktiefond Dalarna	50.37
(18A)	Lannebo Sverige Plus	62.23	(18B)	SEB Stiftelsefond Balanserad	55.37
(19A)	Swedbank Robur Stiftelsefond	61.98	(19B)	Enter Sverige	55.54
(20A)	Swedbank Robur Talenten Aktiefond MEGA	61.97	(20B)	Handelsbanken Japan Tema Criteria	49.08
(21A)	Catella Sverige Aktiefond Hållbarhet	61.81	(21B)	Carnegie Rysslandsfond	47.39
(22A)	AMF Aktiefond Sverige	61.74	(22B)	Swedbank Robur Sverigefond Mega	59.24
(23A)	Länsförsäkringar Sverige Indexnära	61.65	(23B)	Swebank Robur BAS action	55.60
(24A)	Carnegie Sverigefond A	61.64	(24B)	Carnegie Strategifond A	56.10
(25A)	SEB Sverige Expanderad	61.63	(25B)	SEB Dynamisk Aktiefond	56.53
(26A)	Swedbank Robur Ethica Global MEGA	61.58	(26B)	Swedbank Robur Sverigefond	59.22
(27A)	Swedbank Robur Ethica Global	61.53	(27B)	Skandia USA	52.71
(28A)	Nordea Sverige Passiv Icke Utdeling	61.52	(28B)	Swedbank Robur Global Emerging Markets	49.56
(29A)	Finlandsfond-A1 SEK	61.48	(29B)	SEB WWF Nordenfond	60.03
(30A)	KPA ETISK aktiefond	61.37	(30B)	Catella Småbolagsfond	53.97

Notes:
(i) Trust Ref. 7B and 9B did not have an ESG score.

Table 1b

Top ESG funds matched by fund size and inception date

Fund Ref.	*FUND SIZE (Billion SEK)*	*INCEPTION DATE*	*Fund Ref.*	*FUND SIZE (Billion SEK)*	*INCEPTION DATE*
(1A)	1.44	10/99	(1B)	1.42	03/99
(2A)	3.47	08/05	(2B)	3.37	11/06
* (3A)	3.08	01/94	* (3B)	3.20	4/95
* (4A)	3.99	04/99	* (4B)	3.93	3/00
(5A)	0.63	03/07	(5B)	0.63	01/06
(6A)	2.40	11/99	(6B)	2.34	11/99
(7A)	6.64	12/98	(7B)	6.54	03/99
(8A)	0.65	10/95	(8B)	0.89	04/97
* (9A)	2.36	01/03	* (9B)	2.33	09/00
(10A)	7.98	09/00	(10B)	8.11	08/00
(11A)	2.09	06/90	(11B)	2.14	12/90
* (12A)	9.36	08/00	* (12B)	9.18	03/00
(13A)	4.38	08/00	(13B)	4.28	10/99
* (14A)	16.04	04/89	* (14B)	17.81	04/84
(15A)	22.05	12/98	(15B)	22.45	02/98
(16A)	2.68	05/95	(16B)	2.85	01/98
(17A)	0.053	10/03	(17B)	1.01	03/07
(18A)	7.66	12/08	(18B)	7.45	11/07
* (19A)	0.62	09/00	* (19B)	0.66	11/99
* (20A)	4.13	12/94	* (20B)	4.05	04/95
* (21A)	3.72	02/98	* (21B)	3.75	10/97
(22A)	26.39	12/98	(22B)	25.18	11/95
(23A)	13.65	11/08	(23B)	13.23	03/07
(24A)	15.79	01/87	(24B)	14.49	08/88
(25A)	10.62	11/73	(25B)	11.18	01/77
(26A)	13.77	11/02	(26B)	13.46	10/02
(27A)	3.27	03/80	(27B)	3.36	09/91
(28A)	10.15	09/08	(28B)	9.47	02/07
(29A)	0.38	10/98	(29B)	0.37	01/99
(30A)	6.96	03/99	(30B)	6.76	02/98

Notes:
(i) Trust Ref. is the same as table 1.
(ii) * denotes that those trusts have been matched with regards to investment holdings.
(iii) Fund size by 31st of March 2019.

Table 2a

Monthly returns on Swedish domiciled open-ended mutual funds

Treatment group				*Control group*			
Fund Ref.	Mean	St.Dev.	*No. Obs*	Fund Ref.	Mean	St.Dev.	*No. Obs*
(1A)	0.0054	0.0368	*120*	(1B)	0.0063	0.0253	*120*
(2A)	0.0098	0.0435	*120*	(2B)	0.0109	0.0527	*120*
(3A)	0.0055	0.0361	*120*	(3B)	0.0121	0.0529	*120*
(4A)	0.0056	0.0416	*120*	(4B)	0.0057	0.0363	*120*
(5A)	0.0092	0.0392	*120*	(5B)	0.0082	0.0313	*120*
(6A)	0.0090	0.0422	*120*	(6B)	0.0050	0.0409	*120*
(7A)	0.0059	0.0356	*120*	(7B)	0.0056	0.0171	*120*
(8A)	0.0097	0.0456	*120*	(8B)	0.0063	0.0572	*120*
(9A)	0.0083	0.0472	*120*	(9B)	0.0069	0.0340	*120*
(10A)	0.0089	0.0463	*120*	(10B)	0.0080	0.0276	*120*
(11A)	0.0081	0.0465	*120*	(11B)	0.0097	0.0338	*120*
(12A)	0.0058	0.0361	*120*	(12B)	0.0085	0.0354	*120*
(13A)	0.0114	0.0483	*120*	(13B)	0.0112	0.0344	*120*
(14A)	0.0091	0.0420	*120*	(14B)	0.0038	0.0548	*120*
(15A)	0.0103	0.0437	*120*	(15B)	0.0079	0.0338	*120*
(16A)	0.0056	0.0386	*120*	(16B)	0.0111	0.0526	*120*
(17A)	0.0091	0.0471	*120*	(17B)	0.0117	0.0450	*120*
(18A)	0.0128	0.0508	*120*	(18B)	0.0052	0.0271	*120*
(19A)	0.0037	0.0182	*120*	(19B)	0.0092	0.0448	*120*
(20A)	0.0092	0.0368	*120*	(20B)	0.0041	0.0431	*120*
(21A)	0.0097	0.0449	*120*	(21B)	0.0085	0.0653	*120*
(22A)	0.0105	0.0445	*120*	(22B)	0.0107	0.0470	*120*
(23A)	0.0097	0.0444	*120*	(23B)	0.0054	0.0267	*120*
(24A)	0.0106	0.0426	*120*	(24B)	0.0089	0.0336	*120*
(25A)	0.0092	0.0413	*120*	(25B)	0.0085	0.0344	*120*
(26A)	0.0086	0.0315	*120*	(26B)	0.0096	0.0477	*120*
(27A)	0.0098	0.0372	*120*	(27B)	0.0107	0.0354	*120*
(28A)	0.0096	0.0454	*120*	(28B)	0.0060	0.0471	*120*
(29A)	0.0084	0.0480	*120*	(29B)	0.0088	0.0413	*120*
(30A)	0.0082	0.0345	*120*	(30B)	0.0106	0.0448	*120*
Average	***0.0086***	***0.0412***		***Average***	***0.0082***	***0.0401***	
SIX RX	***0.0109***	***0.0039***		***SIX RX***	***0.0109***	***0.0039***	

Table 2b

Sharpe Ratio, Treynor Ratio, Jensen´s Alpha, and beta comparison

Treatment group					*Control group*				
Fund Ref.	Sharpe Ratio	Treynor Ratio	Jensen's alpha (%)	β_P	Fund Ref.	Sharpe Ratio	Treynor Ratio	Jensen's alpha (%)	β_P
(1A)	0.474	0.083	- 0.043 *	0.729	(1B)	0.816	0.141	- 0.003**	0.505
(2A)	0.778	0.116	- 0.001 *	1.013	(2B)	0.727	0.115	(- 0.000)	1.159
(3A)	0.494	0.089	- 0.004**	0.692	(3B)	0.377	0.121	- 0.002	0.570
(4A)	0.435	0.079	- 0.004**	0.792	(4B)	0.508	0.089	- 0.004 *	0.718
(5A)	0.804	0.131	- 0.001	0.835	(5B)	0.884	0.195	(-0.000)	0.490
(6A)	0.729	0.110	- 0.002 *	0.970	(6B)	0.393	0.138	- 0.003	0.401
(7A)	0.544	0.097	- 0.004**	0.692	(7B)	1.062	0.179	- 0.003 *	0.350
(8A)	0.729	0.118	- 0.001	0.981	(8B)	0.365	0.086	- 0.003	0.841
(9A)	0.598	0.090	- 0.003 *	1.081	(9B)	0.676	0.109	- 0.003**	0.725
(10A)	0.660	0.099	- 0.002 *	1.065	(10B)	0.978	0.149	- 0.002	0.626
(11A)	0.590	0.089	- 0.003 *	1.065	(11B)	0.992	0.262	0.001	0.445
(12A)	0.525	0.090	- 0.004 *	0.730	(12B)	0.814	0.342	0.001	0.292
(13A)	0.825	0.130	(0.000)	1.073	(13B)	1.139	0.283	0.003	0.484
(14A)	0.747	0.118	- 0.001	0.925	(14B)	0.210	0.066	- 0.004	0.602
(15A)	0.815	0.123	- 0.001	1.009	(15B)	0.795	0.155	- 0.001	0.597
(16A)	0.473	0.081	- 0.004 *	0.781	(16B)	0.740	0.117	(- 0.000)	1.159
(17A)	0.662	0.114	- 0.001	0.951	(17B)	0.908	0.145	0.001	0.984
(18A)	0.891	0.141	0.002	1.120	(18B)	0.613	0.094	- 0.004 *	0.611
(19A)	0.622	0.099	- 0.005	0.397	(19B)	0.705	0.107	- 0.002**	1.021
(20A)	0.857	0.133	- 0.001 *	0.826	(20B)	0.297	0.115	- 0.004	0.383
(21A)	0.743	0.113	- 0.001	1.024	(21B)	0.445	0.119	(-0.000)	0.844
(22A)	0.818	0.124	(-0.000)	1.025	(22B)	0.794	0.120	(-0.000)	1.083
(23A)	0.755	0.114	- 0.001	1.021	(23B)	0.651	0.101	- 0.004 *	0.595
(24A)	0.864	0.132	- 0.001	0.975	(24B)	0.903	0.140	- 0.001	0.750
(25A)	0.760	0.115	- 0.001**	0.947	(25B)	0.843	0.167	- 0.001	0.601
(26A)	0.927	0.177	- 0.001	0.573	(26B)	0.699	0.105	- 0.002**	1.102
(27A)	0.902	0.146	(- 0.000)	0.800	(27B)	1.050	0.266	0.002	0.487
(28A)	0.727	0.110	- 0.001	1.040	(28B)	0.420	0.091	- 0.003	0.752
(29A)	0.599	0.105	- 0.002	0.945	(29B)	0.730	0.114	- 0.002	0.913
(30A)	0.804	0.125	- 0.002	0.769	(30B)	0.823	0.136	(0.000)	0.941
average	**0.705**	**0.113**		**0.895**	***average***	**0.712**	**0.145**		**0.701**

Notes:
(i) **, * denotes significance at the 0.10, 0.05 level or better.
(ii) Parentheses indicates value with more than 3 decimals.

Table 3a

T-test of difference between the pair

Fund Ref.		Fund Ref	T-value 2009 - 2018	T-value 2015 - 2018
(1A)	*vs*	(1B)	*- 0.2590*	- 0.3275
(2A)	*vs*	(2B)	- 0.9110	0.3890
(3A)	*vs*	(3B)	- 1.3932	- 0.6159
(4A)	*vs*	(4B)	0.0919	- 0.1387
(5A)	*vs*	(5B)	0.4118	- 1.9377
(6A)	*vs*	(6B)	0.9397	- 1.2429
(7A)	*vs*	(7B)	0.3619	0.1755
(8A)	*vs*	(8B)	0.6801	- 0.3449
(9A)	*vs*	(9B)	0.8400	- 0.5505
(10A)	vs	(10B)	0.7634	- 0.1362
(11A)	*vs*	(11B)	- 0.3151	- 0.5242
(12A)	*vs*	(12B)	- 0.7927	- 0.2586
(13A)	*vs*	(13B)	0.1807	- 0.4470
(14A)	*vs*	(14B)	1.0130	0.2478
(15A)	*vs*	(15B)	0.9827	- 0.4380
(16A)	*vs*	(16B)	- 2.2057*	- 0.0254
(17A)	*vs*	(17B)	- 0.9432	- 1.8017
(18A)	*vs*	(18B)	3.1416*	1.2885
(19A)	vs	(19B)	- 2.1598*	- 0.7403
(20A)	vs	(20B)	1.2593	- 0.0804
(21A)	*vs*	(21B)	0.0124	- 1.4762
(22A)	*vs*	(22B)	- 0.4650	- 1.1766
(23A)	*vs*	(23B)	2.2922*	0.3205
(24A)	*vs*	(24B)	1.4054	0.2360
(25A)	*vs*	(25B)	0.3514	- 0.6893
(26A)	*vs*	(26B)	- 0.5573	0.5620
(27A)	*vs*	(27B)	- 0.3869	- 0.2936
(28A)	*vs*	(28B)	0.9643	0.0377
(29A)	*vs*	(29B)	- 0.0557	0.6566
(30A)	*vs*	(30B)	- 1.1749	- 0.4274

Notes:
(i) * denotes significance at the 0.05 level or better.

Table 4a

Sharpe ratio

Variable	Obs	*mean*	*Std. Errors*	*Std.Dev*	[95% Conf. Interval]	
Treatment group	**30**	**.7049628**	**.0257141**	**.1408418**	**.6523715**	**.757554**
Control group	**30**	**.7118188**	**.0443471**	**.242899**	**.6211189**	**.8025188**
Diff	**30**	**-.0068561**	**.050528**	**.2767535**	**-.1101975**	**.0964854**
Mean(diff) = mean (**ESG-NESG**)					**t**	**= -0.1357**
					Degrees of freedom	**= 29**
Ha: Mean (diff) <0	Ha: Mean (diff) !=0			Ha: Mean (diff) > 0		
Pr (T < t) = **0.4465**	Pr (|T| < | t|) = **0.8930**			Pr (T > t) = **0.5535**		

Table 4b

Treynor ratio

Variable	Obs	*mean*	*Std. Errors*	*Std.Dev*	[95% Conf. Interval]	
Treatment group	**30**	**.1129304**	**.003986**	**.0218325**	**.0104778**	**.1210828**
Control group	**30**	**.145495**	**.0117778**	**.0645098**	**.1214066**	**.1695834**
Diff	**30**	**-.0325646**	**.0125491**	**.0687341**	**-.0582304**	**-.0068989**
Mean(diff) = mean (**Treatment group -Control group**)					**t**	**= -2.5950**
					Degrees of freedom	**= 29**
Ha: Mean (diff) <0	Ha: Mean (diff) !=0			Ha: Mean (diff) > 0		
Pr (T < t) = **0.0073**	Pr (|T| < | t|) = **0.0147**			Pr (T > t) = **0.9927**		

Table 5a

Test for correlation

	Return	*ESG-high*	*Risk premium*	*Jensen's alpha*	*Size*	*Financial service*	*Manufacturing*	*Healthcare*	*Technology*	*Energy*	*Financial service x ESG high*	*Healthcare x ESG high*
Return	**1.0000**											
ESG-high	0.0071	**1.0000**										
Risk premium	0.9700	0.0068	**1.0000**									
Jensen's alpha	0.0412	- 0.15818	0.0399	**1.0000**								
Size	0.0032	- 0.0430	0.0031	0.0590	**1.0000**							
Financial service	- 0.0194	0.1400	- 0.0188	- 0.2155	- 0.0807	**1.0000**						
Manufacturing	0.0203	0.1690	0.0197	0.1185	- 0.0483	- 0.4971	**1.0000**					
Healthcare	0.0061	- 0.0000	- 0.0059	0.1634	- 0.1837	- 0.0780	- 0.2197	**1.0000**				
Technology	- 0.0044	- 0.2154	- 0.0043	0.0362	0.1673	- 0.1990	- 0.5606	- 0.0880	**1.0000**			
Energy	- 0.0105	- 0.1302	- 0.0102	- 0.0826	0.0534	- 0.0547	- 0.1540	- 0.0242	- 0.0617	**1.0000**		
Financial service x ESG high	- 0.0160	0.3333	- 0.0155	- 0.3536	0.0279	0.7935	- 0.3944	- 0.0619	- 0.1579	- 0.0434	**1.0000**	
Healthcare x ESG high	0.0049	0.1302	0.0047	0.0361	- 0.1922	- 0.0547	- 0.1540	0.7011	- 0.0617	- 0.0169	- 0.0434	**1.0000**

Table 5b

Correlation	*Return*	*Return$_{t-1}$*
Return	**1.0000**	
Return$_{t-1}$	0.0567	**1.0000**

Table 5c

Variable	*VIF*	*1/VIF*
Sharpe ratio	30.82	.032450
Treynor ratio	20.95	.047742
ESGhigh	15.40	.064917
Manufacturing	12.40	.079997
Manufacturing x ESG high	12.14	.082377
Financial service x ESG high	6.43	.155441
Technology	4.76	.209987
Financial service	4.34	.230349
Healtcare	3.12	.320016
Healthcare x ESG high	2.60	.383943
Size	2.49	.402332
Jensen's alpha	2.35	.424795
Energy	1.65	.604596
Risk premium	1.01	.994698
Mean VIF	8.61	

Table 5d

Variable	*VIF*	*1/VIF*
Financial service x ESG high	3.94	.253888
Manufacturing	3.14	.317278
Financial service	3.03	.329886
ESG high	2.80	.357771
Size	2.32	.430456
Healthcare x ESG High	2.13	.469519
Healthcare	2.03	.492577
Jensen's alpha	1.89	.527805
Technology	1.62	.616924
Energy	1.08	.922273
Risk premium	1.00	.995041
Mean VIF	2.27	