UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Investing in Sustainable Stocks

An Empirical Evaluation of Large Public Companies in Sweden

Bachelor of Science (B.Sc.) in Finance

Andreas Bengtsson & Jakob Johnson

Spring 2018

Supervisor: Jon Williamsson

Abstract

This thesis investigates if investing in a sustainable index yields higher risk-adjusted returns than investing in an ordinary index. Return data from the Swedish stock index OMXS30 was collected, spanning the period 2006-2017. Based on the OMXS30, a sustainability index called ESGS10 was constructed manually and then used to compare risk-adjusted returns with OMXS30. The Sharpe Ratio, Sortino Ratio, Reward-to-VaR and Reward-to-CVaR are used as risk-adjusted return measures in order to include both standard risk and tail risk. The return distributions of the indices are shown to be non-normal and leptokurtic as well as exhibiting positive skewness. We find evidence that investing in ESGS10 produces higher risk-adjusted returns than the investing in the OMXS30, although the differences between the indices are not significant on a 5% level.

Keywords

Sustainability, Investing, Risk-Adjusted Returns, Standard Risk, Tail Risk

Abbreviations

ESGS10 **—** Manually constructed sustainability index
OMXS30 **—** An index consisting of the 30 most traded stocks on Nasdaq Stockholm
RAR **—** Risk-adjusted returns

Table of Contents

1 Introduction

1.1 Background

One could argue a keyword over the last few years has been sustainability. This trend and focus has started spilling over into financial sectors, and can now be a very important variable for investment valuations and decision making, according to Vesty (2011). This will lead to fundamental changes in financial markets, and investors will have to ask themselves whether to align their investment strategy with this trend or choose to stand still.

Sustainable investing or the more widely known term socially responsible investing (SRI) can be defined as an investment approach that includes environmental, social, and governance (ESG) aspects for portfolio selection and strategic investment decisions (Global Sustainable Investment Alliance, 2017). SRI has grown at a rapid rate in recent time. At the start of 2016 SRI investment assets globally were estimated to be $22.9 trillion, having risen from $13.3 trillion at the outset of 2012, according to numbers from Global Sustainable Investments (GSI).

As a concept, SRI investing has been the subject of much discussion, where two different prevailing sentiments have developed. The first view, that SRI leads to lower profitability and higher risk, has its foundation in the well-known Markowitz Portfolio Selection paper from 1952. This paper provided the groundwork for the opponents of the argument that SRI can provide increased risk-adjusted returns relative to a more traditional portfolio selection. (Charles, Darne, and Fouilloux, 2015a) The main argument used for this position is that the Markowitz Portfolio Theory concluded that investors can achieve the highest risk-adjusted returns when they have greater potential for diversification and selection. Actively screening certain sectors or companies implies decreasing your spectrum of choice and should therefore also lead to lower risk-adjusted returns (Markowitz, 1952).

The implication of SRI investing most often involve some sort of screening, and hence a smaller investment universe. As above explained, this smaller investment universe can be used to argue that SRI investing therefore experience a loss of portfolio efficiency. Additionally, even if sustainable companies outperform non-sustainable companies on average, evidence of cost to SRI can still be profound, since if only a single non-sustainable company outperforms the sustainable company the investor suffers an implicit cost (Adler and Kritzman, 2008).

The second view argues that excluding (i.e. negative screening) non-sustainable companies, make the remaining companies more investment-worthy. This is due to the fact that the excluded companies are seen as not viable in the long-term due to its unsustainable nature (RBC Global Asset Management, 2016). Those who position themselves to hedge against more long-term exogenous and physical risk will likely be able to generate excess returns. Examples of such risks are climate disasters and increased regulations (Lee & Moscardi, 2017). Furthermore, a paper by El Ghoul, Guedhami, Kwow and Mishra (2011) found support for the notion of socially responsible practices correlating with higher valuations and lower cost of equity. This can mainly be attributed to improving responsible employee relations and environmental policies, resulting in reputational improvements.

1.2 Purpose and Research Question

Previous research has yielded mixed results regarding the potential abnormal returns of SRI and with this in mind, further research is necessary. Consequently, the purpose of this thesis is to evaluate whether a sustainable index can deliver higher risk-adjusted returns on the stock market than a comparable ordinary index. This will be done specifically for large public companies in Sweden, due to a lack of current research in this area and because large companies are likely more sensitive to ESG factors, due to reputational concerns and public opinion, among others (Arvidsson, 2010). In this thesis a sustainable index is defined as an index consisting of companies with the highest ESG-scores. The sustainable index is proxied by the manually constructed index ESGS10. The ordinary index is proxied by OMXS30 which is a selection of the 30 most traded stocks on Nasdaq Stockholm.

Does investing in ESGS10 generate higher risk-adjusted returns than investing in OMXS30?

1.3 Disposition

The paper starts with a more general topic introduction and concept background before subsequently stating the research question and the purpose of this paper. The ensuing chapter, Theoretical Framework, serves as a more thorough presentation of previous research and theory on matters relating to our research, as well as presenting the risk-adjusted return measures we use. The Data and Methodology chapter highlights the methodology used for the data collection process, in conjunction with a discussion regarding the data characteristics, namely potential limitations that might decrease external and internal validity. This chapter also describes more in-depth how the risk-adjusted return measures, that we use for answering the research question, and subsequent statistical tests are performed and calculated.

The results are shown in the chapter Empirical Results and Analysis. This is where results concerning risk-adjusted return measures, return distributions, and statistical significance are first presented, followed by a more thorough analysis as to, among else, how these numbers ought to be interpreted and what conclusions can be drawn as a consequence. Lastly, in the final chapter, Conclusion, the research question stated earlier is answered and the thesis is concluded and brought to a close by stating the main takeaways.

2 Theoretical Framework

2.1 Previous Research

The research area concerning SRI and its potential for abnormal returns often yield contradicting results and conclusions. It is important to keep in mind that research papers, including those presented below, differ by region, data type, and choice of return measure.

Klassen and McLaughlin (1996) investigated the link between environmental management and stock prices. Good environmental management was proxied by environmental awards, while poor environmental management was proxied by environmental crises. The sample consisted of 140 publicly traded companies in the U.S. during the 1985-1991 period. They hypothesized that strong environmental management result in higher future stock prices. The study showed that abnormal positive returns were linked to good environmental management. De and Clayman (2015a) found similar results when they investigated the connection between ESG-scores and stock price performance. They restricted their sample to the U.S. and gathered data from the Thomson Reuters Corporate Responsible Dataset between 2007-2012 and found three interesting results. Firstly, there was a non-significant positive correlation between high ESG-scores and abnormal positive stock returns. Secondly, results were enhanced upon the exclusion of the companies with the lowest ESG-scores. Lastly, in addition to the abovementioned results, they found a strong and significant negative correlation between ESG-scores and volatility.

A paper by Charles et al (2015b) compared sustainable stock indices to equivalent ordinary stock indices, on a regional as well as global scale. The data time frame was 2005-2010. They used several proxies for risk in their calculations, ranging from standard risk to tail risk. The findings showed that even though a number of ESG-indices displayed higher risk-adjusted returns and lower volatility, other ESG-indices displayed opposite results, contributing to inconclusive and contradictory results. Lean and Nguyen (2014) conducted a similar study between 2004-2013 and found that the four sustainable indices used in the study produced lower risk-adjusted returns than their ordinary index counterparts.

Rennings, Schröder and Ziegler (2003) sought to establish how environmental and social performance affect stock prices. Their sample consisted of 300 European companies spanning the period 1996-2001. An econometric analysis was conducted which revealed a significant, positive relationship between stock prices and the environmental performance of the industry in which the company in question operated. The social performance however, had a negative effect. On top of the aforementioned results, the relative environmental and social performance of the company compared to its peers, had no effect.

A study was conducted by Dorfleitner and Halbritter (2015) regarding the effect of ESG-scores on financial performance. Their sample consisted of 4209 American companies included in the ASSET4 ESG-rating universe spanning the period 2002-2011. Furthermore, additional ESG-scores from KLD and Bloomberg were collected to include variations in how different ESG-score providers measure sustainability. Carhart's four-factor regression model was used to investigate the relationship between ESG-scores and excess returns. However, no significant relationship was found, questioning the validity of an investment strategy based on ESG.

2.2 Risk-Adjusted Return Measures

In order to measure the risk-adjusted returns of ESGS10 and OMXS30, several different measures of risk-adjusted return will be utilized as suggested by Charles et al (2015c). The measures to be used are the *Sharpe Ratio*, *Sortino Ratio*, *Reward-to-VaR*, and the *Reward-to-CVaR.*

2.2.1 Standard Risk

The *Sharpe Ratio* (SR) was introduced by William F. Sharpe in 1966 and is a measure of excess return in relation to volatility, defined as the standard deviation of the returns. Such a measure allows for a more comprehensive view of the asset rate of return (Berk and DeMarzo, 2014). The SR has arguably become the most popular measure of risk-adjusted returns and is used by a myriad of different investors and institutions worldwide (Weisman, 2002a).

The excess return is defined as follows:

$$Excess\ Return = I_{R,i} - R_f = ð_{R,i} \qquad \text{(I)}$$

Where $I_{R,i}$, R_f, and $ð_{R,i}$ is defined as the return for index i, the risk-free interest rate, and excess return for index I, respectively.

The SR is defined as follows:

$$Sharpe\ Ratio = \frac{ð_{R,i}}{\sigma_i} \qquad \text{(II)}$$

Where $ð_{R,i}$ and σ_i are defined as the excess return for index i and the standard deviation of the returns for index i, respectively.

A potential downside with the SR is the assumption of a Gaussian return distribution. If the return distribution is non-normal this can cause the SR to yield misleading results (Pav, 2016a).

The *Sortino Ratio* (SOR), first introduced in an article by Price and Sortino in 1994, offers an alternative to the SR. Unlike the SR, the SOR defines risk as exclusively the downside deviation of returns. Only deviation for returns below the mean return is accounted for. (Charles et al, 2015d). With this adjustment in mind, the SOR could be more preferable as a measure of risk-adjusted return, since investors most likely do not mind positive return deviation. Such positive return is desirable, not objectionable (Jin, Markowitz and Zhou, 2006).

The SOR is defined as follows:

$$Sortino\ Ratio = \frac{ð_{R,i}}{\sigma_{dd,i}} \qquad \text{(III)}$$

Where $ð_{R,i}$ and $\sigma_{dd,i}$ are defined as the excess return for index i and the downside-deviation of the returns for index i, resprectively.

2.2.2 Tail Risk

The first risk-adjusted return using tail risk is *Reward-to-VaR* (R-to-VaR) which is a modification of the SR where standard deviation is substituted for Value-at-Risk (VaR) as the measurement of risk (Dowd, 2000). VaR is designed to measure the size of the maximum loss over a given time period within a specific confidence interval (Jorion, 2002). Examples of how to interpret and derive VaR values are given in Figure A.

Figure A

How to interpret and derive VaR values

The question asked in VaR**: What loss level are we X% confident will not be surpassed in N business days?**

Example 1) If you calculate a VaR value of $10 000 for an investment over one day with a confidence level of 99%, it means that you can with 99% likelihood say that your losses will not exceed $10 000 the next day.

Example 2) A project has a 98% chance of $1M in profit, 1,5% chance of a loss of $1M, and 0,5% chance of a loss of $5M. The VaR at confidence level 99% for this project is then a loss of $1M.

VaR measures extreme risk since it answers a simple question of how bad things can get, due to the measurement of potential extreme losses found in distribution tails (Gustafsson and Lundberg, 2009). There are numerous advantageous features of this tail risk measure, the most prominent that it gives you a single number and is fairly straightforward for interpretation, in addition to capturing a vital aspect of risk. However, VaR calculated using the variance-covariance method can give inaccurate values when applied to return distributions that feature significant positive or negative kurtosis, because of the assumption of distribution normality. The depiction of potential losses at the extreme ends of distributions is then not accurate since it is tied to the chosen interval (Yamai and Yoshiba, 2002). Additional drawbacks of the variance-covariance method all originate from the estimation process. If variance and covariance inputs are incorrect, error terms and tail densities could be extorted. Furthermore, because of the non-stationary nature of the variables, an inability to deal with variable changes over time might be cause for concern. (NYU Stern, 2018a)

The variance-covariance method will be employed in this thesis, but not exclusively. Another VaR computation method called the historical simulation method is also to be used extensively. The historical simulation method is similar to the variance-covariance method, but as the name suggests, there is no need for estimation. Variance, covariance and other inputs are derived from the actual distribution. Potential causes for concern regarding the variance-covariance method might be offset to a certain extent by this method, but there are still some deficiencies associated to this method. Since it uses past data to determine future risk exposure, caution is advised since the past is not always a prelude of the future. Moreover, for assets which there is no historical data available, such as new market assets, this method is of no use. (NYU Stern, 2018b)

Despite the potential limitations of VaR, it has been incorporated into the practices of regulatory bodies because of its positive contribution in highlighting extreme risk. In 1995 the Basel Committee of Banking Supervision adopted VaR as a means to quantify risk and it was the basis for calculating capital requirements for commercial banks (Jorion, 1996). In 1996, the Securities and Exchange Commission listed VaR as the one of three ways large publicly listed companies in the United States can measure market risk (Koch, 2006).

VaR is defined as follows:

$$Value\ at\ Risk = \mu_i + \sigma N^{-1}(X) \tag{IV}$$

Where μ_i, σ, and $N^{-1}(X)$ are defined as the mean of the returns from index i, standard deviation of the returns for index I, and the inverse cumulative distribution, respectively.

The definition is the formula for the variance-covariance method and it assumes normal distribution. The other method, historical simulation, is calculated by determining loss levels at different percentiles. The thesis will employ both methods, but the VaR values using the latter method will be used in R-to-VaR, due to the abundance of historical data available and the lack of need for making potentially incorrect assumptions such as distribution normality.

R-to-VaR is defined as follows:

$$Reward - to - VaR\ Ratio = \frac{ð_{R,i}}{|VaR|} \tag{V}$$

Where $ð_{R,i}$ is defined as the excess return for index i.

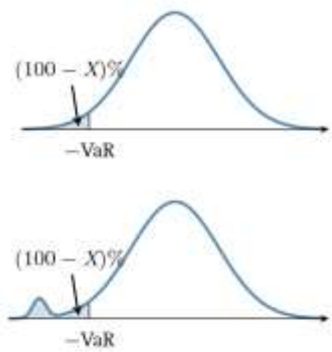
As mentioned previously, the denominator $|VaR|$ is the historical simulation VaR, defined as an absolute value.

The second risk-adjusted return using tail risk is *Reward-to-CVaR* (R-to-CVaR) which substitutes the denominator term in SR for Conditional Value-at-Risk (CVaR). CVaR is an evolution of VaR and gives a value of expected loss, conditional on the loss being greater than the VaR level (Rockafeller and Uryasov, 2002).

CVaR is more coherent then VaR, since it captures tail distribution and density in a more holistic manner. A distribution can have the same VaR, but drastically different CVaR, as illustrated in Figure B, where the return distribution graph below distribution has a much higher CVaR.

Figure B

Distributions with the same VaR but different CVaR



Source: Wimmer, Risk Management and Financial Institutions University of Mannheim

CVaR is defined as follows:

$$CVaR = \mu_i + \sigma_i \frac{e^{-(N^{-1}(X))^2/_2}}{\sqrt{2\pi}(1-X)} \qquad \text{(VI)}$$

Where μ_i, σ_i, and $N^{-1}(X)$ are defined as the mean of the returns from index i, standard deviation of the returns for index i and the inverse cumulative distribution, respectively.

Similar to the VaR computations, CVaR is computed in this thesis using two different methods, one based on an assumption of normal distribution (Formula VI) and another based on the actual historical data. CVaR gives a more accurate description of tail-risk in return distributions which feature kurtosis, since it takes into account all extreme losses. This property is a reason why the Basel Committee on Banking Supervision in 2016 will move away from the use of VaR in favour of CVaR (Basel Committee of Banking Supervision, 2016). CVaR also fulfils the sub-additivity property of risk, which the VaR does not (Tasche, 2002). With all this in mind, the R-to-CVaR can therefore be seen as a more thorough measure of risk. The $|CVaR|$ to be included in R-to-CVaR is based on the actual distribution, similar to the denominator $|VaR|$ in R-to-VaR.

R-to-CVaR is defined as follows:

$$Reward - to - CVaR = \frac{ð_{R,i}}{|CVaR|} \qquad \text{(VII)}$$

Where $ð_{R,i}$ is defined as the excess return for index i.

3 Data and Methodology

An evaluation and comparison between the two indices will be carried out, using multiple measures of risk-adjusted return. Risk-adjusted returns are to be used because it will provide a more comprehensive view of stock performance due to the inclusion of risk.

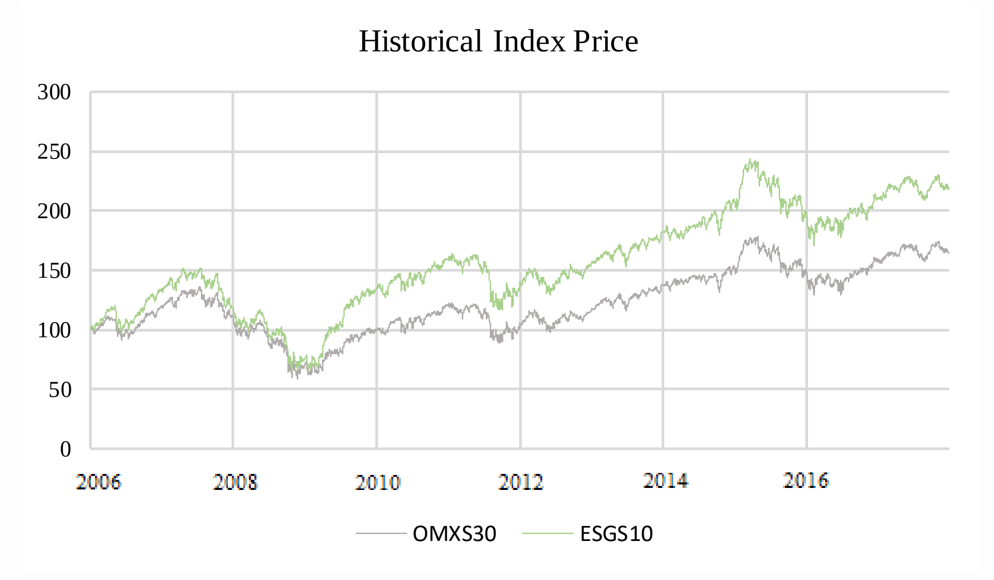
3.1 Measurement Considerations

An important aspect to take into consideration regarding risk-adjusted returns is the measure of choice and more specifically how risk is defined. The Sharpe ratio has become a very popular tool for evaluating risk-adjusted return according to Sharpe (1994) and Weisman (2002b). However, this measure has some deficiencies, none more so than that it measures risk with volatility, therefore making no distinction between upside and downside variability of returns. In order to account for this deficiency, this thesis will use not only standard deviation as the standard risk measure, but also a measure using downside deviation exclusively, as advocated by Price and Sortino (1994). Increasingly though, it is important that one take into consideration risk relating to tail distributions, and not assuming a classical Gaussian distribution. Research by Mandelbrot (1963a) and Fama (1963) challenged these Gaussian return distributions, highlighting the need for using different methods in order to account for tail risk in return distributions (Pazarbosi, 2013). In light of the variety of ways to view risk, both tail risk measures such as Value-at-Risk and Conditional Value-at-Risk, in addition to standard risk measures such as standard deviation and downside deviation, will be used to determine risk-adjusted returns.

3.2 Data Gathering

The data for this thesis is derived from two indices, each of which is viewed as a single security or asset. The OMXS30 index is comprised of the 30 most traded stocks on the Swedish stock market, Nasdaq Stockholm. The second index, ESGS10, constructed manually specifically for this thesis, consists of the 10 most sustainable companies from those in the OMXS30 index. The level of sustainability for a company is based on the Thomson Reuters ESG Score.

Figure C



Horizontal axis: Year
Vertical axis: Index price

All price data relevant for the two indices was collected from Thomson Reuters for a period spanning from January 1st 2006 to December 31st 2017. Because most daily returns are close to zero[1], the choice of whether or not to logarithmize returns is slightly inconsequential (Miskolczi, 2017). It is plausible to assume the returns reflect a stochastic process, since the price of a stock have a tendency to follow a Brownian motion (Shi, 2010). For a non-stochastic process, it can be preferable to use logarithmic returns, see Hudson and Gregoriou (2010). Due to the high plausibility of a stochastic process in the return series, and because logarithmic returns are inferior to simple returns when calculating changes in wealth over a period, the choice to use non-logarithmic returns seems well-founded, although there are some drawbacks associated with it (Hudson and Gregoriou, 2010)

3.3 ESGS10

In order to evaluate whether a sustainable index can generate higher risk-adjusted returns than an ordinary index, a manually constructed index was created. This manually constructed index, ESGS10, is comprised of the ten most sustainable stocks of those in the OMXS30 index. To create such an index requires, of course, quantifying sustainability, despite the potential problems associated with that quantification because of the subjective nature of the term sustainability and lack of an objective measure.

[1] See Figure G and Figure H in Appendix.

The numerical metric we chose to act as a sustainability proxy, and serve as our sustainability quantifier, was the Thomson Reuters ESG Score, as previously used by De and Clayman (2015b). This measure was the most thorough and comprehensive since the score is the result of 178 different metrics and ten different categories, and was therefore deemed the best alternative. The different weights of the categories are provided in Figure I.[2] By using ESG-score as the only positive screening variable, this should provide an adequate indication of whether sustainable investing generate excess returns in relation to ordinary investing.

Table 1

Companies	ESG Score*	Weight**
LM Ericsson	87,60	10,67%
AstraZeneca	86,99	10,60%
ABB	86,78	10,57%
Nordea Bank	82,38	10,03%
Swedbank	81,09	9,88%
Electrolux	80,46	9,80%
H&M AB	79,81	9,72%
Telia Company	79,35	9,67%
Boliden	78,99	9,62%
SKF	77,53	9,44%
SUM	**820,98**	**100,00%**

*Scores taken from Thomson Reuters Eikon ESG Score.

**The ESG score of the single company divided by the sum of all ten ESG scores.

As shown in the table above, the scores provided by Thomson Reuters allow for the companies to be ranked, according to their level of sustainability, determined by the ESG-score. In order to create the ESGS10 index, the ESG weight of each company in the ESGS10 index is calculated. That is then used to weigh each company's return in accordance with their ESG weight to arrive at the aggregated return for ESGS10 index.

3.4 Limitations

There are some limitations with the data used in this thesis. We assume constant ESG-scores for the relevant time period. However, whether or not this is an actual limitation could be contested, seeing that ESG-scores likely do not trend downwards, but upwards. An assumption of constant ESG-scores might therefore be quite a neutral assumption.

[2] See Appendix.

Additionally, the positive screening of the 10 most sustainable companies that make up ESGS10 is not of all the 30 companies on OMXS30, but rather a positive screening of 29 companies on the OMXS30 index. This complication is to due there being no satisfactory aggregated ESG-score or similar data on the company Essity AB, hence the exclusion of Essity AB from the positive screening process to select the ten companies that are part of the sustainable ESGS10 index.

A potential concern that ESGS10 consists of only 10 companies and OMXS30 of 30 companies, and the asymmetrical impact on variance for each as a result, is legitimate and something to have in mind. This concern can be partially offset by the fact that the two indices are highly correlated and that based on total market capitalization, ESGS10 account for roughly 53.33% of OMXS30.

Further limitations include the rather small population of each index and the non-representativeness (only larger companies are researched) of each population when results might be attributed to an entire stock market. The potential downside of a smaller population can be viewed in the prism of a trade-off for getting more accurate numbers determining sustainability levels since they only exist for larger companies. Lastly but not least, since larger companies likely have a higher sensitivity towards reputational concerns and public opinion than smaller companies, it is important to bear this in mind when drawing conclusions.

3.5 Risk-Adjusted Return Measures Computations

3.5.1 Sharpe Ratio and Sortino Ratio

The SR and SOR are both calculated on a yearly basis, over the period January 1st 2006 - December 31st 2017. The excess return for each ratio is calculated using daily returns for each year to determine the yearly return and afterwards subtracting the risk-free rate. Due to the exclusive focus on the Swedish stock market, the selected appropriate risk-free rate proxy to be used was the reference rate provided by the Swedish Central Bank.[3] For each relevant year, an annual average of this rate is calculated to be used as part of the yearly excess return calculations. As explained by Riksbanken (2018) and Hanqvist (2013), the reference rate is

[3] The Swedish Central Bank refers to Riksbanken and the reference rate refers to “referensränta”, itself provided by the Swedish Central Bank.

used a benchmark for commercial interest rates as well as government bonds, hence strengthening the argument for it to be used as a proxy for the risk-free rate.

Once the excess return is calculated, the denominator in the formula needs to be determined. For the Sharpe Ratio, this is fairly straightforward. A yearly standard deviation is calculated, by first calculating the daily standard deviation for a given year, and then multiplied with the square root of the number of trading days. This methodology is then performed for each year to get a yearly Sharpe Ratio for both indices.

The calculation of a Sortino Ratio has many similarities to that of a Sharpe Ratio. However, as previously shown in Formula III the denominator term is different, as only downside deviation is to be included. The downside deviation is computed by calculating the standard deviation of the returns below the average daily return for a given year, and then multiplying the daily downside deviation with the square root of the number of trading days for the given year. All SR and SOR computations are performed in Excel.

Additionally, test for significant differences in the risk-adjusted return (using standard risk) results between the two indices, OMXS30 and ESGS10, were performed in Stata.

3.5.2 Reward-to-VaR and Reward-to-CVaR

The risk-adjusted returns using standard deviation and downside deviation, respectively, described in 3.4.1, were both measurements for risk-adjusted return where *standard risk* acted as the proxy for risk. For both R-to-VaR and R-to-CVaR, *tail risk* is used as the proxy for risk. The *tail risk* used in each measure is perfectly hinted in the names. For R-to-VaR, Value-at-Risk is the risk proxy and for R-to-CVaR, Conditional Value-at-Risk serve as the risk proxy.

Two methods each are used to determine both the VaR and CVaR, namely the variance-covariance method, see Formula IV for VaR and Formula VI for CVaR, and the historical simulation method. The former is based on an assumption of normal distribution, whilst the latter is based on the actual return distribution. VaR and CVaR are calculated in Excel for each year, using both methods.

The R-to-VaR and R-to-CVaR consist of the excess return, see 3.4.1 for how it is derived, in the numerator and the absolute value of VaR or CVaR in the denominator. An example of a computation is given in Table 2 to the right. A higher R-to-VaR or R-to-CVaR indicate a better tail-risk adjusted return, and vice versa.

Table 2

OMXS30	
Year	*2017*
Annual Return	3,28
Risk-Free Rate	-0,50
5% VaR	-1,025
Reward-to-VaR	**3,688**

$$\frac{3{,}28-(-0{,}5)}{1{,}025}=3{,}688$$

If there are large disparities between the VaR when using the variance-covariance method versus using the historical simulation method, one might be able to use this information to say something about tail characteristics and corresponding tail density, since the former method assumes normal distribution. The same reasoning applies for CVaR. Because of this aspect, the value in performing both methods when calculating VaR and also CVaR, is quite cogent. Statistical tests are therefore carried out to test the significance between VaR values using the variance-covariance method and using the historical simulation method. The same test and procedure is also conducted for CVaR values.

3.6 Statistical Tests

3.6.1 Test of Normality

The distribution of data is of importance whenever statistical analysis is to be conducted. This thesis will compute risk-adjusted return measures such as the SR which makes an assumption about the normality of the underlying population (Pav, 2016b). Therefore, tests of normality will be applied to the dataset in order to ascertain if the returns for OMXS30 and ESGS10 follow a normal distribution. The thesis will perform the *Jarque-Bera* test and *Shapiro-Francia* test in Stata to analyse the return distribution of the dataset.

The *Jarque-Bera* test, also known as the *Bowman-Shenton* test was outlined in an article by Bowman and Shenton (1975). The test employs the distribution metrics skewness and kurtosis in order to estimate the distribution of the data. A dataset which is normally distributed has a skewness coefficient of zero and kurtosis coefficient of three.[4] The *Jarque-Bera* test sets up the joint null hypothesis that the dataset in question have a skewness and kurtosis coefficient equal to zero. The test then calculates the JB test statistic that follows the Chi-Square distribution. The test statistic can be used to reject, or not to reject the null hypothesis, given the chosen level of significance (Jaggia and Kelly, 2013a).

The Jarque-Bera JB test statistic is defined as:

$$JB = \frac{n}{6}\left[S^2 + \frac{K^2}{4}\right] \qquad \text{(VIII)}$$

Where n, S and K is defined as the sample size, skewness coefficient and kurtosis coefficient, respectively.

The *Shapiro-Francia* test, first introduced in an article by Shapiro and Francia (1972), is a modification of the *Shapiro-Wilk* test. The main difference between the tests is that *Shapiro-Francia* is valid for larger sample sizes. The *Shapiro-Wilk* test is defined for $4 \leq n \leq 2000$ observations, while *Shapiro-Francia* allows for $5 \leq n \leq 5000$ observations (Stata, 2018). This property makes it suitable for the dataset of this thesis. Similar to the *Shapiro-Wilk* test, the *Shapiro-Francia* sets up the null hypothesis that the sample comes from a normal distribution, and then compares the order statistics from the normal distribution with the ordered values from the sample. A W statistic is then computed, which can be used to reject, or not reject the null hypothesis, depending on the level of significance (Shapiro and Wilk, 1965).

The Shapiro-Francia W statistic is defined as:

$$W = \frac{\left(\sum_{j=1}^{n} w_j x_{(j)}\right)^2}{(n-1)s^2} \qquad \text{(IX)}$$

w_j – Function of the mean variance and covariance of the order statistic.

$x_{(j)}$ – The largest of n observations.

s^2 – The unbiased estimate of the population variance.

n – Sample size.

[4] Neutral kurtosis can be zero as well, it depends on whether the formula includes a subtraction of three or not.

3.6.2 Test of Risk-Adjusted Return Measures

When conducting statistical tests one can choose either parametric tests or non-parametric tests. Parametric tests are characterised by the assumption that the population being tested follow a certain distribution, such as the normal distribution. Furthermore, they require a large sample size and are concerned with the testing of population means. The t-test is a classic example of a parametric test (Jaggia and Kelly, 2013b).

If the data sample does not meet the standards parametric tests set out, non-parametric or distribution free tests can be applied instead. As the name suggests, such tests do not assume a particular distribution for the data. They are also still applicable for small sample sizes. Non-parametric tests are valid for data of ordinal or nominal scale, as opposed to parametric tests who demand interval or ratio scale. The main downside with non-parametric tests is the potential loss of power (higher likelihood of Type II error) when a non-parametric test is chosen, even though a parametric test could have been used (Jaggia and Kelly, 2013c).

When testing for statistically significant differences between the risk-adjusted return of the two indices, we opted for the use of a non-parametric test since our data sample of 24 observations per measure is considered small. Furthermore, financial data tend to follow non-normal distributions which make parametric tests ill suited (Bollerslev, Todorov and Li, 2012). As a result of these considerations, we chose the Wilcoxon rank-sum test for independent samples to analyse the data.

The Wilcoxon rank-sum test in named after Frank Wilcoxon (1892-1945) who introduced it in a *Biometrics Bulletin* article in 1945 (Wilcoxon, 1945). The test is used for determining if the population median differs from a chosen value. To conduct the test all the data is arranged after size, and thereafter assigned a rank. The data is then divided into group X and group Y. The test seeks to make a statement regarding the population of group X, compared to group Y.

The critical values for Wilcoxon rank-sum test statistic W is defined as:

$$z = \frac{W - \mu_W}{\sigma_W} \qquad \text{(X)}$$

Where the test statistic W is the value associated with the minimum of n_1 and n_2, $\min(n_1, n_2)$:

a) If $n_1 \leq n_2$, then $W = W_1$ or

b) If $n_1 > n_2$, then $W = W_2$,

where W_1 and W_2 represent the sum of the ranks of the values in both samples.

(Jaggia and Kelly, 2013d)

We conduct the Wilcoxon rank-sum test under the null hypothesis that the population medians of the risk-adjusted return measures calculated for ESGS10 and OMXS30 are equal, thus testing if the difference between them are significant. In addition to the Wilcoxon rank-sum test we calculate a statistic we call the Probability of Sustainability Surplus.[5]

4 Empirical Results and Analysis

4.1 Risk-Adjusted Return Measures

The data in Table 3 and 5 is a compilation of the SR, SOR, R-to-VaR and R-to-CVaR for ESGS10 and OMXS30 over the period 2006-2017 as well as for each individual year. The metrics should be interpreted in an ordinal sense, as opposed to a cardinal sense, since they are not actual returns, i.e. the change in wealth, for a given period. The numbers in the Table 3 and 5 are important because it can provide rankings and information in order to decide whether one index is better than the other.

[5] See Table 12 in Appendix.

Table 3

Risk-Adjusted Returns Using Standard Risk Measures

	Sharpe Ratio (SR)		Sortino Ratio (SOR)		Sustainability Excess*	
Year	*OMXS30*	*ESGS10*	*OMXS30*	*ESGS10*	*SR*	*SOR*
2017	0,37	0,40	0,66	0,62	0,03	-0,04
2016	0,43	0,33	0,56	0,44	-0,09	-0,12
2015	-0,06	-0,32	-0,09	-0,49	-0,26	-0,40
2014	0,70	0,94	1,08	1,45	0,24	0,36
2013	1,30	1,49	1,86	2,28	0,19	0,41
2012	0,46	0,63	0,65	0,91	0,17	0,26
2011	-0,64	-0,63	-0,91	-0,88	0,01	0,03
2010	1,01	0,96	1,53	1,42	-0,05	-0,10
2009	1,22	3,03	1,92	4,90	1,81	2,98
2008	-1,03	-1,16	-1,75	-1,95	-0,12	-0,20
2007	-0,51	-0,65	-0,71	-0,91	-0,14	-0,19
2006	0,88	1,67	1,18	2,21	0,79	1,03
Average:	***0,34***	***0,56***	***0,50***	***0,83***	***0,21***	***0,33***
Entire Period***	***0,80***	***1,02***	***1,11***	***1,43***	***0,22***	***0,31***

*ESGS10 minus OMXS30, coloured for sustainability excesses.

** SR and SOR for the entire period (2006-2017), calculated as one dataset.

The higher the value, the higher the risk-adjusted return.

Table 4

Descriptive Statistics

Standard RAR Measures

	SR		SOR	
	OMXS30	**ESGS10**	**OMXS30**	**ESGS10**
Mean	0,344	0,557	0,498	0,833
Observations	12	12	12	12
Maximum	1,295	3,033	1,919	4,897
Minimum	-1,035	-1,155	-1,746	-1,945
Standard Deviation	0,757	1,174	1,153	1,830
Kurtosis*	2,053	2,772	2,275	3,194
Skewness	-0,485	0,488	-0,559	0,604
SD Mean**	0,219	0,339	0,333	0,528

*Kurtosis calculation without subtraction of 3, hence value < 3 indicate platykurtic distribution.

** Standard error of mean.

Table 3 indicates that the SR and SOR for ESGS10 outperforms OMXS30 both on a yearly average and when the risk-adjusted returns are computed for the entire period as one whole dataset. These results are in line with that of De and Clayman (2015c) where they found a positive relationship between ESG ratings and share price. However, the higher risk-adjusted returns for ESGS10 relative to OMXS30 are not statistically significant, as later shown in section 4.3. OMXS30 also perform better over the last period, 2015-2017. For both indices, the row Entire Period consistently generate higher values than the yearly average. Furthermore, the SOR produces higher values than the SR, as expected due to the elimination of downside deviation in the computation of SOR.

Table 5

Risk-Adjusted Returns Using Tail Risk Measures

	Reward-to-VaR		**Reward-to-CVaR**		**Reward-to-VaR**		**Reward-to-CVaR**	
	OMXS30	*ESGS10*	*OMXS30*	*ESGS10*	*OMXS30*	*ESGS10*	*OMXS30*	*ESGS10*
Year	*5%**				*1%***			
2017	3,69	3,85	3,08	2,96	2,65	2,48	2,20	2,05
2016	4,56	3,99	2,75	2,14	2,23	1,62	1,43	1,12
2015	-0,51	-2,83	-0,43	-2,22	-0,40	-1,75	-0,31	-1,56
2014	6,79	9,73	5,22	7,14	4,25	6,20	3,58	5,45
2013	11,69	17,55	8,80	10,83	7,18	7,56	6,19	6,54
2012	4,13	6,35	3,07	4,31	2,43	3,27	2,14	3,03
2011	-5,80	-5,46	-4,17	-4,02	-3,67	-3,51	-3,07	-2,93
2010	8,66	8,88	7,24	6,97	6,22	6,14	6,00	8,67
2009	11,74	30,40	9,20	23,25	7,39	19,74	7,09	27,98
2008	-9,91	-10,60	-7,39	-8,40	-6,68	-7,06	-6,14	-8,63
2007	-4,06	-5,25	-3,28	-4,26	-2,79	-3,63	-2,74	-1,87
2006	9,89	17,33	5,79	10,97	3,94	8,02	3,75	9,60
Average:	***3,41***	***6,16***	***2,49***	***4,14***	***1,90***	***3,26***	***1,68***	***4,12***
Entire Period***	***27,07***	***35,65***	***18,55***	***23,90***	***15,33***	***19,35***	***12,00***	***15,53***

*VaR and CVaR calculated at 95% level.

**VaR and CVaR calculated at 99% level.

*** Reward-to-VaR and Reward-to-CVaR for the entire period (2006-2017), calculated as one dataset.

The higher the value, the higher the risk-adjusted return.

Coloured indicate higher (more positive or less negative) values for ESGS10.

Table 6

Descriptive Statistics

Tail RAR Measures

	R-to-VaR				R-to-CVaR			
	5%		1%		5%		1%	
	OMXS30	**ESGS10**	**OMXS30**	**ESGS10**	**OMXS30**	**ESGS10**	**OMXS30**	**ESGS10**
Mean	3,406	6,162	1,896	3,257	2,490	4,138	1,676	4,121
Observations	12	12	12	12	12	12	12	12
Maximum	11,740	30,396	7,389	19,739	9,202	23,252	7,089	27,981
Minimum	-9,906	-10,600	-6,675	-7,064	-7,395	-8,398	-6,141	-8,631
Standard Deviation	7,098	11,674	4,463	7,122	5,312	8,602	4,088	9,158
Kurtosis*	2,115	2,660	2,238	3,513	2,117	3,144	2,233	5,030
Skewness	-0,534	0,516	-0,535	0,749	-0,482	0,638	-0,461	1,360
SD Mean**	2,049	3,370	1,288	2,056	1,533	2,483	1,180	2,644

*Kurtosis calculation without subtraction of 3, hence value < 3 indicate platykurtic distribution.

** Standard error of mean,

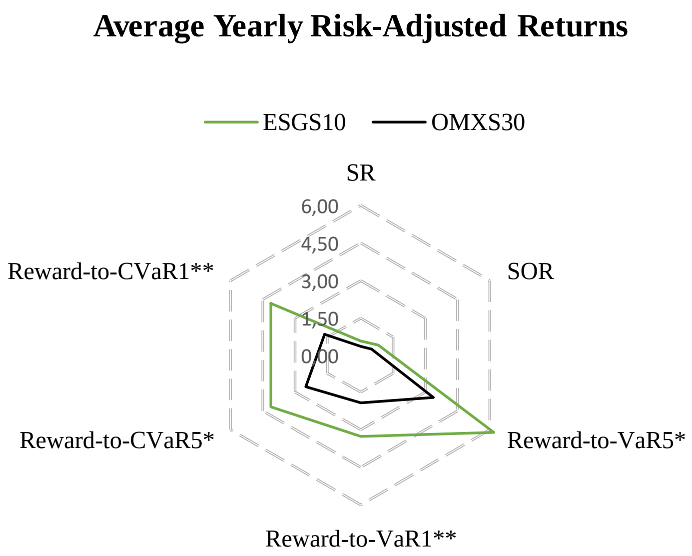
Similar to Table 3, Table 5 shows superior risk-adjusted returns for ESGS10 on a consistent basis, although similar to the standard risk-adjusted return measures, the risk-adjusted return difference between the two indices is not statistically significant. A closer inspection of the risk-adjusted returns, calculated using tail risk, bring about the noteworthy observation, similar to Table 3, that over the last three years, 2015-2017, the values for OMXS30 almost routinely have higher scores, with one exception. The period preceding and following the 2008 financial crisis show more positive values for ESGS10, whilst the year of the financial crisis itself show higher values for OMXS30. This trend could indicate something with regards to a higher multiplying factor effect of the more sustainable index in periods of drastic price movements. In periods of high volatility and more drastic price movement, the ESGS10 performs even better when there is an upward trend and even worse when there is a downward trend.

From Table 4 and 6, a notable observation is the higher variability of risk-adjusted returns for ESGS10, which is clearly evident judging by the higher standard deviation as well as the greater range between minimum and maximum values. Even though nominal returns displayed lower volatility for said index, the volatility of risk-adjusted returns display the opposite. As previously mentioned, this might imply a higher multiplying factor effect of some sort for ESGS10 when sustainable returns are adjusted for risk. The extent of such a phenomenon is not further examined in this paper. One should keep in mind, however, that ESGS10 still generated higher average risk-adjusted returns, as well as positive skewness.

The presence of positive skewness for the risk-adjusted returns of ESGS10 is attractive for more risk-averse investors (Francis, 1975).

In Tables 3 and 5, risk-adjusted returns for EGS10 are consistently higher than for OMXS30. Both the yearly averages of risk-adjusted returns and the risk-adjusted returns for the whole dataset calculated as one indicate higher risk-adjusted returns for the more sustainable index than for the ordinary. These results are in line with an overall macro trend of more sustainable investing. In the past, socially responsible investing amounted to the use of mainly negative screening, i.e. the exclusion of entire sectors that are deemed as sustainability offenders, often termed "sin stocks", or a more norm based negative screening where individual companies are excluded on the basis of environmental damage or human rights violations (Colle and York, 2008). Over the last few years, positive screening strategies or ESG integration for investment analysis and decision making has become more prevalent and underline the overall trend. The existence of empirical evidence supporting the notion of higher financial returns and lower risk for responsible investing is rather convincing, albeit often encountering difficulties of finding significance (Zhang, 2017a). This thesis delivers similar results, as seen in the summary of risk-adjusted returns in Figure D and Table 9.

Figure D



* Calculated at 5% level.

** Calculated at 1% level.

There are many potential explanations for the higher risk-adjusted returns for the more sustainable index. To start with, one should differentiate between causality, reverse causality, and correlation. Just because more sustainable companies might perform better, culminating in higher stock prices, does not necessarily mean ESG is a positive causal factor on better performance. It could simply be that better performing companies can devote and engage more resources toward improving ESG factors, implying reverse causality instead. Assuming the former, rationales for higher financial returns in SRI compared to ordinary investing point to the emergence of sustainable investing over the last decade, therefore increasing demand on these types of investments, both from institutional investors to smaller investors. Increasing regulations and the use of soft law and power on legislator and governing levels, have served to increase awareness and move public- and private sector investments to actively invest in a more sustainable manner.

An implication of a higher ESG-score for companies might be the implicit correlation to a more long-term perspective and outlook by firm management. This interaction between a higher degree of company sustainability and decreasing agency costs, reputation enhancement, better attraction of human capital, and a more well-run firm in a general sense result in better firm performance, and thus higher financial valuations, as mentioned by Charlo, Moya and Munoz (2017a), Ferrell, Liang and Renneboog (2016a), and Tsoutsoura (2004a). The higher financial returns for investing in such companies over the last decade should come as a consequence of a more accurate inclusion of the beforementioned factors in firm valuations.

Looking ahead, socially responsible investing is likely to become mainstream and replace the current view of conventional investing, resulting in potentially mitigating abnormal financial returns for sustainable investing. Be that as it may, this is if one continues using the same definition and methodology for sustainable investing as the current one. Those who act on the forefront and constantly develop their strategies can still take advantage of a financial return sustainability excess, even though the more conventional fund will increase its ESG score. (Zhang, 2017b)

4.2 Return Distributions and Normality

Speculative financial returns have been shown to have fat-tailed distributions, as well as other prominent features such as volatility clustering and non-normality in the form of leptokurtic distributions, see e.g. Huang and Lin (2004) and Mandelbrot (1963b). A useful tool to detect fat tails in distributions is the presence of positive (or negative) kurtosis. A symbolic definition of kurtosis is the ratio of the average of the fourth power of the deviations from the mean, to the square of the variance (Chissom, 2017). Positive kurtosis indicate "fatter" tails and a higher peak, which might seem like contradictory claims at first. An alternative and helpful way of looking at the tail of a leptokurtic distribution, is to think of it in terms of tail density, i.e. the area in the tail, instead of a more graphical (vertical) height. In this way, longer and skinnier tails can also be classed as fat tails for leptokurtic distributions. A more leptokurtic distribution can therefore increase the likelihood of ending up in the tails of the distribution, whether or not that is preferable depends on the risk aversion of each party. Kurtosis is the same independent of the variance of the distribution (DeCarlo, 1997).

Figure E

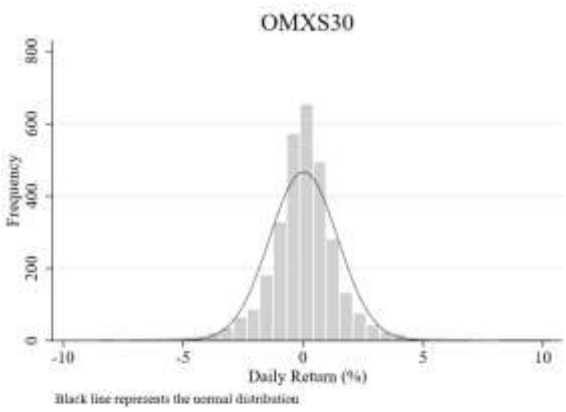
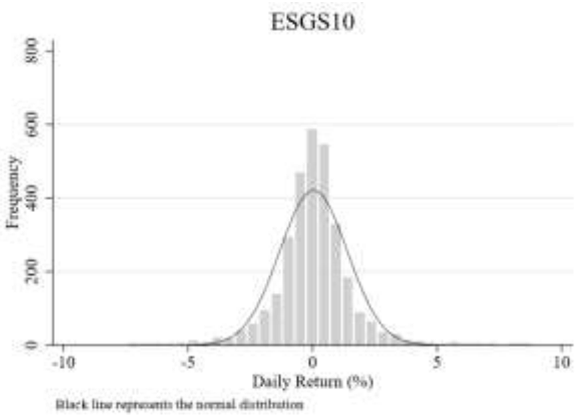


Figure F



Leptokurtic distribution → $k > 0$

Mesokurtic distribution → $k = 0$

Platykurtic distribution → $k < 0$

where k = kurtosis

The return distributions for indices OMXS30 and ESGS10 are leptokurtic, meaning they have excess kurtosis resulting in peakedness and a higher tail density. This is consistent with the results that Charles et al (2015e) found in their study of sustainable stock indices.

This distribution is, as aforementioned, commonplace for financial returns and for symmetrical unimodal distributions.

Table 7

Normality Tests of the Nominal Return Distributions

	OMXS30	**ESGS10**
	2006-2017*	
Shapiro – Francia (prob > z)	0,00001	0,00001
Jarque – Bera (prob > chi^2)	0,0000	0,0000
Skewness	0,1204	0,1321
Kurtosis*	8,0586	7,6138
Average Return	0,0265	0,0352
Standard Deviation**	1,4230	1,3615

*A value above 3 implies high density tails.

*A value below 3 implies low density tails.

*A value equal to 3 implies mesokurtic distribution.

**Daily estimate (%)

*** Values calculated for the whole period, as one dataset.

Table 7 showcases the results from tests of normality applied to the return data for OMXS30 and ESGS10 over the entire period 2006-2017. The tests conducted are the Shapiro-Francia and Jarque-Bera. Moreover, the skewness, kurtosis, average return and standard deviation for the period is displayed.

OMXS30 and ESGS10 exhibit slight skewness to the right side of the return distribution, suggesting it is not completely symmetrical. ESGS10 has a slightly higher degree of skewness to the right as compared to OMXS30. These results go against those found in a study by Hong and Stein (2003) which say that stock returns tend to suffer from negative skewness. ESGS10 exhibits a higher average return over the period 2006-2017. These results are aligned with those of Gray, Kreander, and Power (n.d.) who investigated 40 ethical funds across Europe and found that the majority of the ethical funds performed better than their ordinary benchmark indices.

Both indices show a high degree of kurtosis compared to a normal distribution. The results from the Jarque-Bera test confirms this as the results are highly significant with very low p-values. The results from Shapiro-Francia further confirms the non-normality of the return distributions as it shows very low p-values as well. Similar results, expressly asymmetrical return distributions, was found by Peiro (1999) in his study of eight international stock indices. This evidence that the return distributions for OMXS30 and ESGS10 exhibit asymmetry could potentially present a problem when calculating the SR since it abides by the assumption of a normal return distribution. However, these concerns are alleviated by the results from a study by Eling (2008), which indicate that the SR is still valid despite a non-normal return distribution.

ESGS10 outperforms the OMXS30 both in terms of average return and standard deviation when measured over the entire period. Moreover, the return distribution for ESGS10 is less leptokurtic than the OMXS30 indicating that ESGS10 have less extreme observations at the ends of the tails, as shown in Table 7 and 8. This is consistent with the lower volatility of the ESGS10. These are interesting results since it points towards ESGS10 producing higher returns, at a lower volatility than the ordinary index. The lower volatility of ESGS10 is matched in a study by Boutin-Dufresne and Savaria (2004) who also found that socially responsible companies exhibited lower volatility.

Table 8

All values provided in the table are given in percentages, for example -5 = -5%.

Entire Period (2006-2017)			**CVaR**	**VaR**
ESGS10	*5%*	AD*	-3,19	-2,14
		ND**	-3,22	-2,20
	1%	AD*	-4,90	-3,94
		ND**	-6,09	-3,13
OMXS30	*5%*	AD*	-3,36	-2,30
		ND**	-3,38	-2,31
	1%	AD*	-5,20	-4,07
		ND**	-3,38	-3,28

*Calculated with the Actual Distribution (AD) of returns for each dataset.

**Calculated assuming Normal Distribution (ND) of returns for each dataset.

4.3 Statistical Significance

The Wilcoxon rank-sum test, performed with a null hypothesis of samples coming from populations with equal medians, was utilized to determine whether the difference between the two indices and corresponding risk-adjusted returns was significant. A parametric test, the two-sided t-test, was performed which yielded notably lower p-values than the Wilcoxon rank-sum test. However, based on the discussion in section 3.6.2 a non-parametric is better suited for our dataset. As a consequence, only results from the Wilcoxon rank-sum test are provided.

Table 9

Risk-Adjusted Return Measures Tests of Significance

Wilcoxon rank-sum

	2006-2017	
Sharpe Ratio (SR)*	0,817	
Sortino Ratio (SOR)*	0,773	
	5%	***1%***
Reward-to-VaR**	0,729	0,729
Reward-to-CVaR**	0,817	0,644

*Values shown are p-values (Prob > |z|), and tests of significant difference between the medians of OMXS30 and ESGS10.

**Same as above, but for Reward-to-VaR and Reward-to-CVaR each at both 5% and 1%.

One can see that the p-values from the tests are non-significant. This means that the difference between risk-adjusted returns for EGS10 and OMXS30 is not significant. The insignificant differences found between the risk-adjusted returns are aligned with results from other studies such as those by Gregory, Matatko and Luther (1997a) and Bauer, Koedijk and Otten (2006a). Moreover, it would be unexpected if the results from our thesis were significant due to the origin of ESGS10. ESGS10 was derived from the OMXS30 which mean they both have companies in common and a high degree of correlation as a result. Furthermore, the sample sizes for the indices are small, which reduces the power of the tests (Jaggia and Kelly, 2017e).

5 Conclusion

This thesis sought to establish if SRI produces abnormal returns for large public companies on the Swedish stock market. A sustainable index, ESGS10, was constructed using the most sustainable large public companies included in the OMXS30. The risk-adjusted returns of ESGS10 is then compared to the risk-adjusted returns of OMXS30. The sustainable index ESGS10 produced superior average yearly risk-adjusted returns as well as higher risk-adjusted returns for the entire period, calculated as one whole dataset, which supports the results found by Zhang (2017c). However, results from statistical tests show that the difference between the indices are not significant on a 5% level. These insignificant results match those of Gregory et al (1997b) and Bauer et al (2006b). The return distributions of the ESGS10 and OMXS30 follow a non-normal distribution, as is commonplace for financial data. Furthermore, both indices have excess kurtosis and are leptokurtic. Both indices also exhibit positive skewness, indicating asymmetric distributions. The ESGS10 generates higher average nominal return and lower volatility, when calculated over the entire period.

We believe the higher risk-adjusted returns for the sustainable index over the time period researched can be explained by a range of factors. Such factors include the more long-term perspective held by management of more sustainable companies, in addition to decreased agency costs, reputational improvements and better attraction of human capital, as suggested by Charlo et al (2017b), Ferrel et al (2016b), and Tsoutsoura (2004b). These factors increase the value proposition and hence firm valuations, leading to a higher stock price. Another explanation could simply be the higher demand for sustainable investments, which would drive prices upwards. It is also important to keep in mind that this study is limited to large companies on the Swedish stock market, and potential external validity problems could result as a consequence, when attributing results of this thesis to non-similar populations.

This study implies there are abnormal returns to be found by investing in sustainable companies. However, the extent to which such abnormal returns are present for SRI is still not entirely clear, seeing that the differences in levels of risk-adjusted returns between the two indices are not statistically significant. Future research is advised to use less correlated indices for comparison, since a high degree of correlation might decrease the potential statistical significance of the results. Moreover, a possible extension of this paper could be to further examine the higher multiplying factor effect of risk-adjusted returns for the more sustainable index in times of drastic price movement and high volatility, such as the period surrounding the 2008 financial crisis.

In summary, the results from this thesis show a tendency for higher risk-adjusted returns when investing in the sustainable index relative to an ordinary index.

6 References

Adler, T. & Kritzman, M., 2008. The Cost of Socially Responsible Investing. *The Journal of Portfolio Management*, Fall, 35(1), pp. 52-56.

Arvidsson, S., 2010. Communication of Corporate Social Responsibility: A Study of the Views of Management Teams in Large Companies. *Journal of Business Ethics,* 96(3), pp. 339-354.

Basel Committee on Banking Supervision, 2016. *Minimum capital requirements for market risk,* Basel: Basel Committee on Banking Supervision.

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

Berk, J. & DeMarzo, P., 2014. *Corporate Finance.* Third Edition ed. Boston: Pearson Education.

Bollerslev, T., Todorov, V. & Li, S. Z., 2013. Jump tails, extreme dependencies, and the distribution of stock returns. *Journal of Econometrics,* Volume 172, pp. 307-324.

Boutin-Dufresne, F. & Savaria, P., 2004. Corporate Social Responsibility and Financial Risk. *The Journal of Investing,* Spring, 13(1), pp. 57-66.

Bowman, K. O. & Shenton, L. R., 1975. Omnibus test contours for departures from normality based on Sqrt(b1) and b(2). *Biometrika*, August, 62(2), pp. 243-250.

Charles, A., Darne, O. & Fouilloux, J., 2015. *ESG Indices Financial Risk: Performance measures based on Value-at-Risk approach,* Nantes: Audencia Nantes, School of Management.

Charlo, M. J., Moya, I. & Munoz, A. M., 2017. Financial Performance of Socially Responsible Firms: The Short- and Long-Term Impact. *Sustainability*, September, Volume 9.

Chissom, B. S., 1970. Interpretation of the Kurtosis Statistic. *The American Statistician*, October, 24(4), pp. 19-22.

Colle, S. d. & York, J. G., 2008. Why Wine is not Glue? The Unresolved Problem of Negative Screening in Socially Responsible Investing. *Journal of Business Ethics,* Volume 85, pp. 83-95.

DeCarlo, L. T., 1997. On the Meaning and Use of Kurtosis. *AmericanPsychologicalAssociation,Inc.*, 2(3), pp. 292-307.

De, I. & Clayman, M. R., 2015. The Benefits of Socially Responsible Investing: An Active Manager's Perspective. *The Journal of Investing,* Winter, 24(4), pp. 49-72.

Dorfleitner, G. & Halbritter, G., 2015. The wages of social responsibility — where are they? A critical review of ESG investing. *Rreview of Financial Economics,* Volume 26, pp. 25-35.

Dowd, K., 2000. Adjusting for risk: An improved Sharpe ratio. *International Rreview of Economics and Finance,* Volume 9, pp. 209-222.

El Ghoul, S., Guedhami, O., C.Y. Kwow, C. & R. Mishra, D., 2011. Does corporate social responsibility affect the cost of capital?. *Journal of Banking & Finance,* 1 February, 35(1), pp. 2388-2406.

Eling, M., 2008. Does the Measure Matter in the Mutual Fund Industry?. *Financial Analysts Journal*, 63(3), pp. 1-13.

Fama, E. F., 1963. Mandelbrot and the Stable Paretian Hypothesis. October, 36(4), pp. 420-429.

Ferrell, A., Liang, H. & Renneboog, L., 2016. Socially responsible firms. *Journal of Financial Economics,* Volume 122, pp. 585-606.

Francia, R. S. & Shapiro, S. S., 1972. An Approximate Analysis of Variance Test for Normality. *Journal of the American Statistical Association*, March, 67(337), pp. 215-216.

Francis, J.K., 1975. Skewness and Investors' Decisions. *The Journal of Financial and Quantitative Analysis*, 10(1), pp. 163-172.

Global Sustainable Investment Alliance , 2016. *Global Sustainable Investment Review 2016,* Washington DC: Global Sustainable Investment Alliance .

Gray, R. H., Kreander, N., Power, D. M. & D., S. C., n.d. *The Financial Performance of European Ethical Funds 1996-1998,* Glasgow: Department of Accountancy and Business Finance, University of Glasgow..

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

Gustafsson, M. & Lundberg, C., 2009. *An empirical evaluation of Value at Risk,* Göteborg: University of Gothenburg, School of Business, Economics & Law.

Hanqvist, D., 2013. Finansmarknadens referensräntor - Särskilt om STIBOR. *Juridisk Publikation,* Volume 1, pp. 81-112.

Hong, H. & Stein, J. C., 2003. Differemces of Opinion, Short-Sales Constraints, and Market Crashes. *The Review of Financial Studies,* Summer, 16(2), pp. 487-525.

Huang, Y. C. & Lin, B.-J., 2004. Value-at-Risk Analysis for Taiwan Stock Index Futures: Fat Tails and Conditional Asymmetries in Return Innovations. *Review of Quantitative Finance and Accounting*, Volume 22, pp. 79-95.

Hudson, R. S. & Gregoriou, A., 2015. Calculating and Comparing Security Returns is harder than you think: A Comparison between Logarithmic and Simple Returns. *International Review of Financial Analysis,* March, Volume 38, pp. 151-162.

Jaggia, K. & Kelly, A., 2013. *Business Statistics - Communicating with Numbers.* New York: McGraw-Hill Irwin.

Jin, H., Markowitz, H. & Zhou, X. Y., 2006. A Note on Semivariance. *Mathematical Finance*, January, 16(1), pp. 53-61.

Jorion, P., 1996. Measuring the Risk in Value at Risk. *Financial Analysts Journal*, December, 52(6), pp. 47-56.

Jorion, P., 2002. How Informative Are Value-at-Risk Disclosures?. *The Accounting Review*, October, 77(4), pp. 911-931.

Klassen, R. D. & McLaughlin, C. P., 1996. The Impact of Environmental Management on Firm Performance. *Management Science*, August, 42(6), pp. 1199-1214.

Koch, S., 2006. Using Value-at-Risk for IS/IT Project and Portfolio Appraisal and Risk Management. *The Electronic Journal Information Systems Evaluation*, 9(1), pp. 1-6.

Lean, H. H. & Nguyen, D. K., 2014. Policy uncertainty and performance characteristics of sustainable investments across regions around the global financial crisis. *Applied Financial Economics*, 24(21), pp. 1367-1373.

Lee, L.-E. & Moscardi, M., 2017. *2017 ESG Trends to Watch*, New York City: MSCI.

Mandelbrot, B., 1963. The variation of certain speculative prices. *The Journal of Business*, October, 36(4), pp. 394-419.

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

Miskolczi, P., 2017. Note on Simple and Logarithmic Return. *Applied Studies in Agribusiness and Commerce*, 11(1-2), pp. 127-136.

Pazarbasi, A., 2013. Tail Risk Literature Review. *Alternative Investment Analyst Review*, 1(4)

Peiro, A., 1999. Skewness in Financial returns. *Journal of Banking & Finance*, Volume 23, pp. 847-862.

Price, L. N. & Sortino, F. A., 1994. Performance Measurement in a Downside Risk Framework. *The Journal of Investing*, Fall, 3(3), pp. 59-64.

RBC Global Asset Management, 2017. *Responsible Investing: The Evolution of Ownership*, Toronto: RBC Global Asset Management.

Rennings, K., Schröder, M. & Ziegler, A., 2007. The effect of environmental and social performance on the stock performance of european corporations. *Environmental and Resource Economics*, 37(4), pp. 661-680.

Riksbanken, 2018, avaiable at https://www.riksbank.se/en-gb/statistics/search-interest--exchange-rates/reference-rate/ (last visited 9 May, 2018).

Rockafellar, R. T. & Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, Volume 26, pp. 1443-1471.

Shapiro, S. S & Wilk, M. B., 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, December, 52(3/4), pp. 591-611.

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

Shi, A., 2010. *Stochastic Processes and their Applications in Financial Pricing*, Washington: University of Washington.

StataCorp LLC, 2018, available at https://www.stata.com/manuals13/rswilk.pdf (last visited 7 May 2018).

Tasche, D., 2002. *Expected Shortfall and Beyond*, Frankfurt: Deutsche Bundesbank.

Tsoutsoura, M., 2004. *Corporate Social Responsibility and Financial Performance,* Berkeley: University of California at Berkeley, Haas School of Business.

Weisman, A. B., 2002. Informationless Investing and Hedge Fund Performance Measurement Bias. Summer, 26(4), pp. 82-91.

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, December, 1(6), pp. 80-83.

Value at Risk (VaR), available at https://www.google.se/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwjRm7H5w5fbAhVGVSwKHc-fCZ4QFggpMAA&url=http%3A%2F%2Fwww.stern.nyu.edu%2F~adamodar%2Fpdfiles%2Fpapers%2FVAR.pdf&usg=AOvVaw0LTI6bYAwGj6Qez4ndGCbT (last visited 10 May, 2018)

Vesty, G., 2011. *The influence and impact of sustainability issues on capital investment decisions,* Melbourne: University of Melbourne.

Yamai, Y. & Toshinao, Y., 2002. Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization. *Monetary and Economic Studies*, January, 20(1), pp. 87-121.

Zhang, L., 2017. *The Financial Return of Responsible Investing*, Utrecht: Sustainable Pension Investments Lab.

7 Appendix

Table 10

All values provided in the table are given in percentages, for example -5 = -5%.

ESGS10

	CVaR						VaR					
	Act. D	**Norm.D**		**Act. D**	**Norm.D**		**Act. D**	**Norm.D**		**Act. D**	**Norm.D**	
	5%		***Divergence****	***1%***		***Divergence****	***5%***		***Divergence****	***1%***		***Divergence****
2017	-1,39	-1,53	0,14	-2,00	-2,89	0,89	-1,06	-1,04	-0,02	-1,65	-1,48	-0,17
2016	-3,16	-3,04	-0,13	-6,03	-5,74	-0,29	-1,70	-2,08	0,38	-4,17	-2,95	-1,22
2015	-2,74	-2,88	0,13	-3,89	-5,40	1,51	-2,15	-1,98	-0,17	-3,47	-2,80	-0,68
2014	-1,68	-1,88	0,20	-2,20	-3,58	1,38	-1,23	-1,27	0,04	-1,93	-1,82	-0,11
2013	-1,69	-1,78	0,09	-2,80	-3,41	0,62	-1,04	-1,20	0,16	-2,42	-1,73	-0,69
2012	-2,75	-2,78	0,03	-3,91	-5,27	1,36	-1,87	-1,89	0,02	-3,62	-2,70	-0,93
2011	-4,42	-4,30	-0,13	-6,07	-8,04	1,97	-3,26	-2,97	-0,29	-5,07	-4,18	-0,90
2010	-2,49	-2,66	0,17	-3,09	-5,07	1,97	-1,95	-1,80	-0,15	-2,83	-2,58	-0,24
2009	-3,72	-4,05	0,33	-4,87	-7,85	2,97	-2,85	-2,71	-0,14	-4,39	-3,93	-0,45
2008	-5,01	-5,65	0,64	-6,56	-10,48	3,93	-3,97	-3,93	-0,03	-5,95	-5,50	-0,46
2007	-2,87	-2,89	0,02	-3,46	-5,40	1,93	-2,33	-2,00	-0,33	-3,37	-2,81	-0,56
2006	-3,03	-2,88	-0,15	-5,01	-5,53	0,52	-1,92	-1,94	0,03	-4,15	-2,80	-1,35
Average	-2,91	-3,02	0,11	-4,16	-5,72	1,57	-2,11	-2,07	-0,04	-3,59	-2,94	-0,65
Entire Period**	-3,19	-3,22	0,04	-4,90	-6,09	1,19	-2,14	-2,20	0,07	-3,94	-3,13	-0,80

*Actual Distribution minus Normal Distribution, and the differences calculated are percentage points.

**Var and CVaR (calculated using normal and actual distributions at both 5% and 1% levels, respectively) for the entire period (2006-2017), calculated as one dataset each.

Table 11

All values provided in the table are given in percentages, for example -5 = -5%.

OMXS30

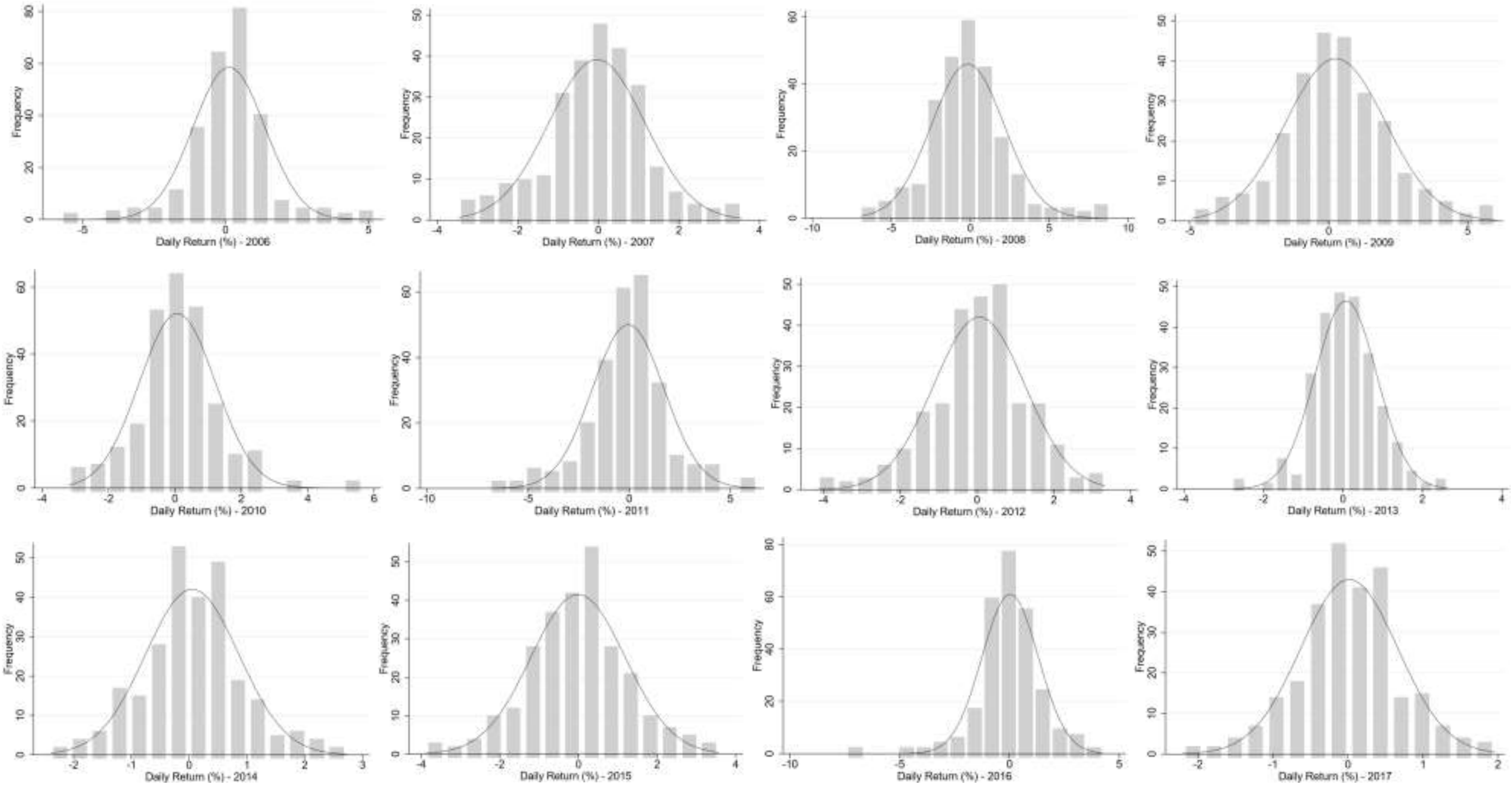
	CVaR						VaR					
	Act. D	**Norm.D**		**Act. D**	**Norm.D**		**Act. D**	**Norm.D**		**Act. D**	**Norm.D**	
	5%		*Divergence**	*1%*		*Divergence**	*5%*		*Divergence**	*1%*		*Divergence**
2017	-1,39	-1,53	0,14	-2,00	-2,89	0,89	-1,03	-1,03	0,01	-1,43	-1,47	0,04
2016	-3,16	-3,04	-0,13	-6,03	-5,74	-0,29	-1,98	-2,17	0,19	-4,06	-3,08	-0,98
2015	-2,74	-2,88	0,13	-3,89	-5,40	1,51	-2,27	-2,06	-0,22	-2,90	-2,91	0,01
2014	-1,68	-1,88	0,20	-2,20	-3,58	1,38	-1,41	-1,38	-0,03	-2,26	-1,97	-0,28
2013	-1,69	-1,78	0,09	-2,80	-3,41	0,62	-1,45	-1,28	-0,17	-2,36	-1,84	-0,52
2012	-2,75	-2,78	0,03	-3,91	-5,27	1,36	-2,13	-1,92	-0,21	-3,62	-2,73	-0,88
2011	-4,42	-4,30	-0,13	-6,07	-8,04	1,97	-3,07	-2,94	-0,13	-4,85	-4,13	-0,71
2010	-2,49	-2,66	0,17	-3,09	-5,07	1,97	-2,24	-1,92	-0,33	-3,12	-2,75	-0,38
2009	-3,72	-4,05	0,33	-4,87	-7,85	2,97	-3,06	-2,88	-0,18	-4,86	-4,14	-0,72
2008	-5,01	-5,65	0,64	-6,56	-10,48	3,93	-4,21	-4,33	0,12	-6,24	-6,06	-0,18
2007	-2,87	-2,89	0,02	-3,46	-5,40	1,93	-2,55	-2,12	-0,43	-3,71	-2,99	-0,72
2006	-3,03	-2,88	-0,15	-5,01	-5,53	0,52	-1,73	-1,93	0,20	-4,34	-2,76	-1,57
Average	-2,91	-3,02	0,11	-4,16	-5,72	1,57	-2,26	-2,16	-0,10	-3,65	-3,07	-0,58
Entire Period**	-3,36	-3,38	0,02	-5,20	-3,38	-1,82	-2,30	-2,31	0,01	-4,07	-3,28	-0,78

*Actual Distribution minus Normal Distribution, and the differences calculated are percentage points.

**Var and CVaR (calculated using normal and actual distributions at both 5% and 1% levels, respectively) for the entire period (2006-2017), calculated as one dataset each.

Figure G

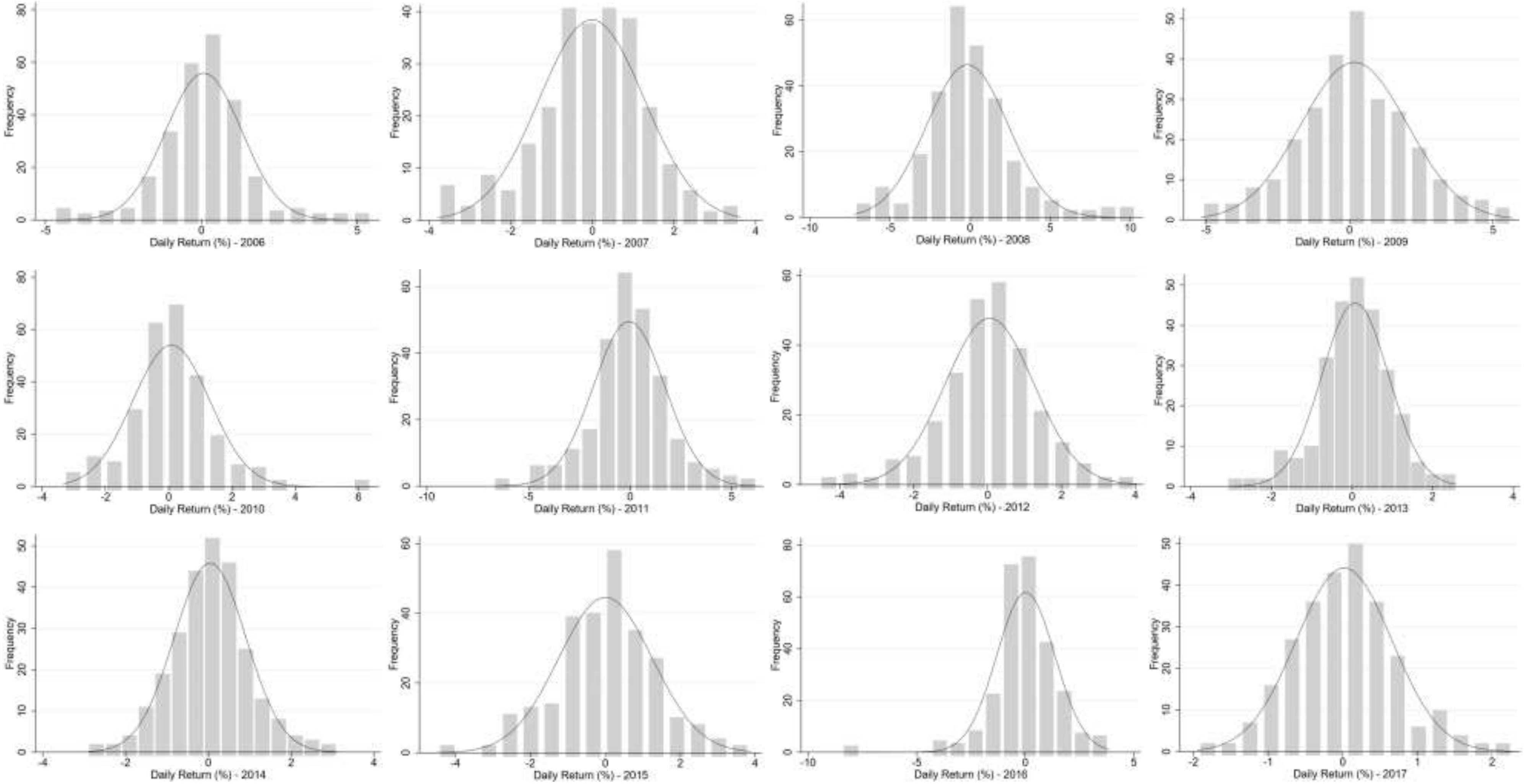
ESGS10



The ESGS10 nominal return distributions for each year. The line represents the normal distribution. As was the case for the entire period, most years show a leptokurtic distribution.

Figure H

OMXS30



The OMXS30 nominal return distributions for each year. The line represents the normal distribution. As was the case for the entire period, most years show a leptokurtic distribution.

Figure I

Weights of ESG Score calculation

Pillar	Category	Weights
Environmental	Resource Use	11%
	Emissions	12%
	Innovation	11%
Social	Workforce	16%
	Human Rights	4,5%
	Community	8%
	Product Responsibility	7%
Governance	Management	19%
	Shareholders	7%
	CSR Strategy	4,5%
TOTAL		100%

source: Thomson Reuters

Table 12

Probability of Sustainability Surplus

	2006-2017	
Sharpe Ratio (SR)	0,528	
Sortino Ratio (SOR)	0,535	
	5%	***1%***
Reward-to-VaR	0,542	0,542
Reward-to-CVaR	0.528	0.556

Note: probability that a randomly selected yearly risk-adjusted return for ESGS10 is larger than a randomly selected yearly risk-adjusted return for OMXS30.

All values above 0,50 indicate a sustainability surplus.