A Study on the Relationship Between a Mutual Fund’s Risk-Adjusted Return and Sustainability:

Do Mutual Funds with High Sustainability Scores Outperform Those with Low Ones?

FRIDA VÄRNLUND

MAX BACCO
A Study on the Relationship Between a Mutual Fund’s Risk-Adjusted Return and Sustainability:

Do Mutual Funds with High Sustainability Scores Outperform Those with Low Ones?

FRIDA VÄRNLUND

MAX BACCO

Degree Projects in Applied Mathematics and Industrial Economics (15 hp)
Degree Programme in Industrial Engineering and Management (300 hp)
KTH Royal Institute of Technology year 2019
Supervisor at Carnegie Fonder, Karl Norin
Supervisors at KTH: Jörgen Säve-Söderbergh
Examiner at KTH: Jörgen Säve-Söderbergh
Abstract

During the past few decades, social responsible investing (SRI) has rapidly grown to become a renowned investment strategy. Because of the contradictory findings on how successful this strategy is in terms of financial return, the aim of this thesis is to compare the performance of sustainable and conventional funds in four different geographical areas during the last three years. With the use of regression analysis, the correlation between the Portfolio Sustainability Score of a fund, which is a Morningstar-provided rating that represents how well a fund incorporates ESG, and its risk-adjusted return is determined.

The final results of this analysis varies among the four geographical regions. The correlation between the two variables is positive in USA and Asia ex-Japan, whereas a negative relationship is found in Europe and the Nordic region. However, the obtained findings are either not of statistical significance or not representative of the given data, implying that there is no difference between the risk-adjusted returns of sustainable versus conventional funds.
Sammanfattning


De slutgiltiga resultaten av denna analys varierar i de fyra geografiska områdena. I USA och Asien där Japan exkluderas är korrelationen positiv medan en negativ korrelation råder i Europa och Norden. Dock är resultaten antingen ickerepresentativa av den givna datamängden eller inte av statistisk signifikans vilket indikerar att det inte är någon skillnad i den riskjusterade avkastningen mellan hållbara och vanliga fonder.
Acknowledgements

We wish to acknowledge our supervisor Jörgen Säve-Söderbergh for his valuable and constructive suggestions throughout the planning and development of this bachelor thesis. Furthermore, we would like to thank Carnegie Fonder for their assistance with the collection of our data as well as insights regarding their implementation of sustainability. Our special thanks are extended to Karl Norin who gave us detailed instructions on how to collect and filter the data, Erik Amcoff who provided us with an elaborate exposition of Carnegie Fonder’s sustainability work and finally Daniel Donatello for his guidance and support.
# Contents

1 Introduction ................................. 7  
  1.1 Background ................................ 7  
  1.2 Purpose .................................. 7  
  1.3 Scope ................................... 8  
  1.4 Problem Formulation ....................... 8  

2 Earlier Studies on the Area ............. 8  

3 Economic Theory ............................. 9  
  3.1 Definition of ESG .......................... 9  
    3.1.1 Environmental ......................... 10  
    3.1.2 Social .................................. 10  
    3.1.3 Governance ............................. 10  
  3.2 Sharpe Ratio .............................. 10  

4 Mathematical Theory ....................... 11  
  4.1 Simple Linear Regression ................... 11  
    4.1.1 Model .................................... 11  
    4.1.2 Assumptions ............................. 11  
    4.1.3 Method of Least-Squares ................. 11  
  4.2 Model Adequacy ............................ 12  
    4.2.1 Residual Scaling ....................... 12  
    4.2.2 Residual Plots .......................... 14  
    4.2.3 Outliers, Leverage and Influential Observations .......................... 15  
    4.2.4 Measure of Influence .................... 15  
    4.2.5 Measure of Model Performance .......... 16  
  4.3 Model Properties ......................... 17  
    4.3.1 Hypothesis Testing .................... 17  
    4.3.2 Confidence Intervals .................... 18  
  4.4 Model Development ......................... 18  
    4.4.1 Coefficient of Determination ............ 18  
    4.4.2 Residual Mean Square ................... 19  
    4.4.3 Transformations ......................... 19  

5 Data ........................................... 19  
  5.1 Delimitations ............................. 19  
  5.2 Data Collection ........................... 20  
    5.2.1 Data Source ............................. 20  
    5.2.2 Sample Selection ....................... 20  
  5.3 Variables ................................. 21  
    5.3.1 Morningstar Portfolio Sustainability Score .......................... 21  
    5.3.2 Risk-Adjusted Return ................... 22  
  5.4 Critics of the Data Set .................... 23  

6 Analysis .................................... 23  
  6.1 Nordic Region ............................. 23  
    6.1.1 Residual Analysis ....................... 23  
    6.1.2 Cook’s Distance .......................... 25  
    6.1.3 CovRatio ................................ 25  
    6.1.4 Model Properties ....................... 26  
    6.1.5 Box Cox ................................. 27  
  6.2 Europe ................................... 27  
    6.2.1 Residual Analysis ....................... 27  
    6.2.2 Cook’s Distance .......................... 29
1 Introduction

1.1 Background

Within the research community, it is well-established that the world stands before numerous challenges due to the effects of global warming. This has led to a vibrant discussion regarding what individuals can and should do to minimize their contribution to the issue. Some of the most common actions the society is urged to take, among others, is to recycle, commute or bike to work as well as eat a more plant-based diet.

However, merely a hundred companies have accounted for seventy percent of the world’s greenhouse gas emissions since 1988 [19]. Therefore, it is clear that actions which individuals can take, similar to the ones mentioned above, have a relatively small effect on the environment compared to the impact companies can have. As a result, new strategies for how individuals can make a significant difference have emerged.

One of the most renowned strategies that allows individuals to influence companies is Social Responsible Investing (SRI). SRI is defined as any investment strategy where the investor considers the environmental, social and governance aspects of a company, but at the same time maintaining its focus on financial return [10]. There are two different approaches to this type of investment strategy. One strategy, known as negative screening, is to exclude companies or industries that do not meet the ethical strategies set by the investor. This could, for example, mean that investors exclude the tobacco, alcohol or weapon industries. The second strategy, positive screening, instead actively selects sustainable firms to invest in.

As a result of the increased popularity in this type of investment strategy, socially responsible investment assets under professional management have grown by more than one-third since 2005 while conventional professionally managed assets have remained flat [7]. This highlights that there is a clear trend in investing in environmental and socially conscious companies within asset management.

1.2 Purpose

At the same time as the trend of social responsible investing has developed, professional fund managers have openly stated that SRI leads to higher financial returns compared to conventional investing [3]. However, despite these statements, multiple reports have concluded otherwise. Some reports claim that social responsible investing does not have an impact on the return [19] and some that it has a negative effect [4]. Therefore, it is clear that the findings on this topic are contradictory and inconsistent.

The purpose of this report is to reach clarifications on this subject by investigating how the sustainability score of a mutual fund is related to its risk-adjusted return. Thereby, the outcome of this paper may be of interest for fund management corporations willing to improve their investing strategies, which is why the project is carried out in cooperation with Carnegie Fonder.
1.3 Scope

To begin with, this thesis is narrowed down to focus on actively managed open-end equity funds. This limitation is to ensure that all sorts of index and exchange traded funds are excluded. Furthermore, this paper investigates four different geographical regions including the Nordic region, Europe, USA and Asia ex-Japan. Japan is excluded from Asia since it is considered to be a developed economy in comparison to the rest of Asia [8]. In each respective region, the correlation between the sustainability score of a fund and its risk-adjusted return is calculated with the use of regression analysis.

1.4 Problem Formulation

The scope of this study can be summarized in the following three questions:

1. How are mutual funds classified in terms of sustainability?
2. What is the potential relationship between the sustainability score of a mutual fund and its risk-adjusted return?
3. How can this potential relationship be implemented by Carnegie Fonder?

2 Earlier Studies on the Area

Studies on social responsible investing (SRI) have been dated back to as early as 1972 [28]. Since then, investments based on environmental, social and governance (ESG) criteria have grown significantly and become an increasing part of asset management. In 2005, already 12% of the money under professional management in the United States was invested according to SRI criteria [4], and this number continues to grow. Globally, there was a 25% increase in the number of assets managed under responsible investment strategies between 2014 and 2016. Furthermore, SRI in Asia (excluding Japan), grew by 16% between 2012 and 2016 [1]. Clearly, SRI is a relevant and intriguing subject due to its fast growth, which has resulted in a large increase in the number of studies on this area.

Therefore, this section addresses the conclusions that earlier studies on the topic of SRI have found, specifically, studies on how sustainable funds perform compared to conventional funds. To get an all-around perspective on the matter, different types of sources are used, including sources from both independent institutions that take an unbiased approach as well as companies that have a more biased view on the topic.

To begin with, a paper that studied the performance and investment style of 103 ethical mutual funds concluded that there is no evidence of a statistically significant difference between ethical and conventional mutual fund returns. More precisely, the study found that ethical mutual funds under-performed in the beginning of the 1990’s due to a so-called catching-up phase. However, over the 1998-2001 period, ethical mutual funds provided average risk-adjusted return matching those of conventional funds due to maturity. Besides models such as the CAPM 1-factor model, this study used more elaborated multi-factor models such as 4-factor asset-pricing model that controls for size, book-to-market and stock price momentum [4].
In addition, other studies that analyzed funds of different countries and time periods have found that the performance of SRI funds is comparable to those of conventional funds. Hence, the restricted investment universe that follows from social screening does not affect the risk-adjusted performance of SRI funds [28].

Furthermore, another paper that used a sample of 42 socially responsible mutual funds, each of which is matched to two randomly selected conventional funds of similar net assets, found no significant difference in investment performance. Moreover, the study found other interesting aspects from the comparison. For example, SRI funds do not differ from conventional funds in degree of portfolio diversification and asset characteristics. Further on, characteristics such as the number of holdings within the fund, the size of the companies and the portfolio concentration were not significantly different between the two groups [5].

Lastly, a meta-study on the subject that aggregated the evidence from more than 2000 individual empirical studies found that 90% of the included studies showed a non-negative relationship between ESG and corporate financial performance [18], supporting the above findings.

In contrast to the results found above, institutions with a more biased perspective on the matter have often concluded that SRI results in higher return. Firstly, a recent study published by Nordea Markets concludes that sustainable firms, in terms of ESG criteria, have higher returns. The study shows that since 2012, firms with high ESG scores have, on average, 5% higher returns compared to firms with low ESG scores [3]. Similarly, Handelsbanken Fonder stated that “sustainability equals return”, suggesting that SRI leads to profitability in the long-term [16]. In addition to these findings, an independent paper investigating how different screening mechanisms affects risk and return reached similar conclusions. This study suggests that positive screening resulted in both higher returns as well as total risk, while the results due to negative screening varied [21].

Evidently, the results from the studies examined above are contradictory. Even though most studies have come to the conclusion that there is a non-negative relationship between sustainability and performance, some studies claim the opposite; namely that high sustainability leads to high returns. It is worth mentioning, however, that the studies concluding the latter have a slightly biased perspective since they have a financial interest in the matter. Therefore, there is convincing evidence that there is minimal difference in risk-adjusted performance between sustainable and conventional funds.

3 Economic Theory

3.1 Definition of ESG

ESG, which stands for environmental, social and governance, are three of the main factors considered when screening a firm’s ethical impact and sustainable practices. ESG investing, also called sustainable investing, impact investing or socially responsible investing (SRI) [9], is a growing approach among investors. Each component of ESG is described in detail below.
3.1.1 Environmental

The first aspect within ESG is the environmental criteria, which oversees a company’s overall impact on the environment. This includes aspects such as energy use, natural resource conservation, pollution and waste. In addition, each company is evaluated regarding which environmental risks might affect the specific company’s income and how well the firm manages these risks. For example, a company might be associated with, and face environmental risks related to its disposal of hazardous waste, its ownership of contaminated land and its compliance with the government’s environmental regulations [9].

3.1.2 Social

The social criteria focuses on a company’s business relationships, and three main aspects within this criteria are evaluated. Firstly, the supply chain of a company is analyzed in order to examine whether each supplier holds the same values that the company claims to stand by. Secondly, the employees’ working conditions are considered and evaluated. Lastly, the stakeholders are identified and whether or not their interests are considered by the company is assessed [9].

3.1.3 Governance

The last component of ESG, governance, looks at the company’s accounting methods and whether they are accurate and transparent. Moreover, it also focuses on the stockholders and whether they are allowed to vote on important issues that will have a great impact on the company. Finally, governance takes into account illegal behavior such as political contributions to obtain favorable treatment [9].

3.2 Sharpe Ratio

William F. Sharpe, an American economist and Nobel Prize winner in Economic Sciences, developed several models to assist with investment decision making. For instance, he constructed what is known as the Sharpe ratio, which is today widely used among investors who seek to maximize their profits with respect to the risk taken. The Sharpe ratio is the average return earned in excess of the risk-free rate per unit of volatility (or total risk). Subtracting the risk-free rate from the mean return enables isolation of the profit associated with the risk-taking activities. Generally, the greater the value of the Sharpe ratio, the more attractive the investment [20]. The Sharpe ratio is defined as

\[
\text{Sharpe ratio} = \frac{R_p - R_f}{\sigma_p}
\]

where

- \( R_p \) = return of portfolio
- \( R_f \) = risk-free rate
- \( \sigma_p \) = standard deviation of the portfolio’s excess return
The Sharpe ratio is mainly used to evaluate a portfolio’s past performance and it explains whether a portfolio’s excess return is a result of too much risk or due to smart investment decisions. Even though a mutual fund might have greater return than its peers, it is only a good investment if those higher returns do not come with additional risk. To summarize, the greater a portfolio’s Sharpe ratio, the better its risk-adjusted performance [20].

There are a few limitations when it comes to using the Sharpe ratio as a financial tool. For instance, the Sharpe ratio makes the assumption that returns are normally distributed, since the standard deviation of returns is used in the denominator. This is, however, not always the case as returns in the financial markets are skewed away from the average due to a large number of unforeseen drops or spikes in prices [20].

4 Mathematical Theory

4.1 Simple Linear Regression

4.1.1 Model

Regression analysis is a commonly used statistical technique for analyzing the relationship between the dependent variable, known as the response variable, and the independent variable(s), often called predictor or regressor variable(s). When there is only one regressor variable involved in the model, it is called a simple linear regression model. This model is mathematically defined as follows:

\[ y = \beta_0 + \beta_1 x + \epsilon \]

where \( \beta_0 \) is the intercept, \( \beta_1 \) is the slope and \( \epsilon \) is a random error component, defined as the difference between the observed value of \( y \) and the straight line \( (\beta_0 + \beta_1 x) \) [22, pg.1-13].

4.1.2 Assumptions

The simple linear regression model has four basic assumptions [22, pg.17-19].

1. Linearity: The response variable \( y \) roughly has a linear relationship with the regressor \( x \).
2. Homoscedasticity: The variance of the residuals is constant across all values of the independent variable.
3. Independent errors: The errors are uncorrelated.
4. Normality: For any value of \( x \), the response variable \( y \) is normally distributed.

4.1.3 Method of Least-Squares

The parameters \( \beta_0 \) and \( \beta_1 \) are unknown and can be estimated using the method of least-squares. This method minimizes the sum of the squares of the differences between the observations \( y_i \) and
the straight line. This gives the least-squares criterion

\[ S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \]

and the least-squares estimators, \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \), must satisfy

\[ \frac{\partial S}{\partial \beta_0} \bigg|_{\hat{\beta}_0, \hat{\beta}_1} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \]

and

\[ \frac{\partial S}{\partial \beta_1} \bigg|_{\hat{\beta}_0, \hat{\beta}_1} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0 \]

Simplifying these two equations gives the least-squares normal equations

\[ n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \]

\[ \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i x_i \]

The solution to the normal equations gives the least-squares estimators of the intercept and slope, \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \). These estimators, in their general form, are given by:

\[ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \]

and

\[ \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \]

where

\[ S_{xy} = \sum_{i=1}^{n} y_i x_i - \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i)}{n} = \sum_{i=1}^{n} y_i (x_i - \bar{x}) \]

and

\[ S_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} = \sum_{i=1}^{n} (x_i - \bar{x})^2 \]

[22, pg.13-15]

4.2 Model Adequacy

4.2.1 Residual Scaling

A residual is defined as the deviation between an observation \( y_i \) and the corresponding fitted value \( \hat{y}_i \). Analyzing and plotting different types of residuals is an effective way to check how well the regression model fits the data and discover model inadequacies. Residuals have zero mean and
their approximate average variance is estimated by
\[ MS_{Res} = \frac{SS_{Res}}{n-p} = \frac{\sum_{i=1}^{n} e_i^2}{n-p} \]

Different types of scaled residuals are described below. These types of residuals are helpful in finding observations that are separated from the rest of the data, i.e. outliers [22, pg.130].

- **Standardized Residuals**
  
  Standardized residuals, given by
  \[ d_i = \frac{e_i}{\sqrt{MS_{Res}}}, i = 1, 2, \ldots, n \]
  have zero mean and approximately unit variance. This scaling is common since the approximate average variance of a residual is estimated by \( MS_{Res} \). As a rule of thumb, standardized residuals with a value \( d_i > 3 \) potentially indicate an outlier [22, pg.130-131].

- **Studentized Residuals**
  
  Studentized residuals, defined as
  \[ r_i = \frac{e_i}{\sqrt{MS_{Res}(1-h_{ii})}}, i = 1, 2, \ldots, n \]
  have constant variance, \( Var(r_i) = 1 \), where \( h_{ii} \) is a measure of the location of the \( i \)th point in \( x \)-space. When the data set is large the variance of the residuals tend to stabilize. Therefore, standardized and studentized residuals may not differ much at all and they generally convey equivalent information. This residual is effective in identifying highly influential points with a large residual and a large \( h_{ii} \) [22, pg.131-133].

- **PRESS Residuals**
  
  Press residuals are defined as \( e_{(i)} = y_i - \hat{y}_i \), where \( \hat{y}_i \) is the fitted value of the \( i \)th response when the \( i \)th point is removed. The residual can also be defined as
  \[ e_{(i)} = \frac{e_i}{1-h_{ii}} \]
  where \( h_{ii} \) is the \( i \)th diagonal element of the hat matrix \( H \). Points with a large \( e_i \) and/or \( h_{ii} \) are often considered influential, and if these values are large it will result in a large PRESS residual, \( e_{(i)} \). Therefore, PRESS residuals are also effective in determining influential points and outliers [22, pg.134].

- **R-Student Residuals**
  
  R-student residuals have almost the same definition as the studentized residuals, the difference being that the variance is estimated by \( S^2_{(i)} \) instead of \( MS_{Res} \). \( S^2_{(i)} \) is an estimation of the variance of a given data set when the \( i \)th observation is removed. The residual is therefore defined as
  \[ t_i = \frac{e_i}{\sqrt{S^2_{(i)}(1-h_{ii})}} \]
where $S^2_{(i)}$ is defined by

$$S^2_{(i)} = \frac{(n - p)MS_{Res} - \frac{e^2_i}{1 - h_{ii}}}{n - p - 1}$$

In many situations, $t_i$ will differ little from the studentized residual $r_i$. However, if the $i$th observation is very influential, then $S^2_{(i)}$ can differ significantly from $MS_{Res}$, causing $t_i$ to greatly differ from $r_i$ as well [22, pg.135].

### 4.2.2 Residual Plots

An effective way to investigate the adequacy of the fit is by analyzing a number of residual plots. The most common plots that should be examined are presented below.

- **Normal Probability Plot**

  The Normal Probability Plot of residuals checks whether the normality assumption for the given data set is true. If the data is normally distributed, the graph will plot approximately a straight line [22, pg.136].

- **Residuals vs. Fitted Values**

  A common residual plot is the Residuals vs. Fitted Values Plot. This plot tests the assumptions of linearity, i.e. whether the relationship between the response and regressor variables is linear and homoscedasticity, i.e. whether there is equal variance along the regression line.

  An optimal Residuals vs. Fitted Values Plot should have a relatively straight red line, showing equal variance along the regression line, and should be distributed symmetrically around the zero-line. Therefore, the plot should resemble a horizontal band where there are no obvious model defects, such as outliers [22, pg.139].

  Figure 1 illustrates common patterns for the Residual vs. Fitted Values Plot. As mentioned above, the ideal shape is plot (a). The second plot, (b), is called an outward-opening funnel pattern which implies that variance increases with increasing values of $y$. Plot (c) indicates a double bow pattern. This scenario generally occurs when $y$ is a proportion between zero and one. Lastly, plot (d), normally referred to as a curved plot, indicates a non-linear relationship [22, pg.139-140].

![Figure 1: Patterns for Residual vs. Fitted Values Plots](image-url)
• **Adjusted-Variable Plots**

Adjusted-Variable Plots, also known as Added-Value Plots or Partial Residual Plots, are helpful in determining whether there exists a relationship between the response and regressor variables. The response variable \( y \) and the regressor \( x_j \) are both regressed against the remaining regressors, and the residuals are then obtained for each regression. By plotting these residuals against each other the plot gives an indication of the marginal relationship between \( y \) and each regressor.

If the response variable and the regressor \( x_j \) have a linear relationship, the plot will show a straight line with a non-zero slope. A perfectly horizontal line (a line with a slope of zero) indicates that there is no linear relationship between the response variable and the regressor under consideration [22, pg.143-144].

4.2.3 **Outliers, Leverage and Influential Observations**

An **outlier**, also called an extreme value, is an observation that is separated from the rest of the data. An outlier is generally detected by noticing that it lies an abnormal distance from other values in the same data set.

A **leverage point** is a point that is far away from the rest of the sample in \( x \)-space, but that still lies on the regression line passing through the rest of the sample. This means that the point has an unusually large \( x \)-value, but still fits the rest of the data. The elements \( h_{ij} \) of the hat matrix \( H \) represents the amount of leverage exerted by the \( i \)th observation \( y_i \) on the \( j \)th fitted value \( \hat{y}_i \).

Lastly, an **influence point** is a point that is unusual both in \( x \)- and \( y \)-space. Such a point is said to pull the model coefficients in a certain direction.

[22, pg.211-212]

4.2.4 **Measure of Influence**

To analyze the level of influence that the different points described above have on a data set, the measures Cook’s Distance, DFBETAS and DFFITS can be used.

• **Cook’s Distance**

As mentioned above, Cook’s distance is a method for measuring influence. Points with a large distance, \( D_i \), have a noticeable influence on the least squares estimator \( \hat{\beta}_i \). This measure is a deletion diagnostic, meaning that it measures the influence of the \( i \)th observation if it is removed from the sample. The distance is generally expressed as

\[
D_i = (M, c) = (\hat{\beta}_{(i)} - \hat{\beta})' M (\hat{\beta}_{(i)} - \hat{\beta}) \sqrt{S^2_{(i)}}(1-h_{ii})
\]

The \( D_i \) statistic can also be rewritten as

\[
D_i = \frac{r_i^2}{p} \frac{h_{ii}}{1-h_{ii}}
\]
It is clear from this formula that $D_i$ is made up of two components; one that reflects how well the model fits the data and one that measures how far $i$th observation is from the rest of the data. If either of these two components are large, it will result in a large value of $D_i$. As a rule of thumb, points for which $D_i > 1$ are considered to be influential points.

[22, pg.215-217]

- **DFBETAS**

$DFBETAS$ is similar to Cook’s distance, a deletion diagnostic. This statistic indicates how much $\hat{\beta}_j$ changes if the $i$th observation were to be deleted. The statistic is given by

$$DFBETAS_{j,i} = \frac{\hat{\beta}_j - \hat{\beta}_{j(i)}}{\sqrt{S^2_{(i)} C_{jj}}}$$

where $C_{jj}$ is the $j$th diagonal element of $(X'X)^{-1}$ and $\hat{\beta}_{j(i)}$ is the $j$th regression coefficient when the $i$th observation is deleted. A large value of $DFBETAS_{j,i}$ indicates that observation $i$ is influential [22, pg.217].

- **DFFITS**

The third measure of influence instead investigates the influence of the $i$th observation on the fitted value. This statistic is defined as

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{\sqrt{S^2_{(i)} h_{ii}}}$$

where $\hat{y}_{(i)}$ is the fitted value of $y_i$ without the $i$th observation. Similarly to the previous statistics, a large $DFFITS_i$ indicates a point with large influence [22, pg.217-218].

### 4.2.5 Measure of Model Performance

In contrast to the diagnostics introduced in the previous section, measures of model performance provide information about the overall precision of estimation [22, pg.219].

- **CovRatio**

$COVRATIO_i$ is a common measure of model performance, defined as

$$COVRATIO_i = \frac{(S^2_{(i)})^p}{MS_{Res}^p} \left( \frac{1}{1 - h_{ii}} \right)$$

Clearly, high leverage will make $COVRATIO_i$ large. Cutoff values are points that are considered influential. The points are cutoff values if $COVRATIO_i > 1 + 3p/n$ or if $COVRATIO_i < 1 - 3p/n$, where $p$ is the number of regressors and $n$ is the number of observations [22, pg.219].
4.3 Model Properties

4.3.1 Hypothesis Testing

Hypothesis testing on the slope and intercept is often used to obtain information regarding the parameters. More importantly, it is used to test the significance of regression, i.e. whether there exists a linear relationship between $x$ and $y$ or not. This procedure assumes that the model errors $\epsilon_i$ are normally and independently distributed with mean 0 and variance $\sigma^2$ [22, pg.22]. The hypotheses to test if the slope equals a constant, say $\beta_{10}$, are

$$H_0 : \beta_1 = \beta_{10}, \quad H_1 : \beta_1 \neq \beta_{10}$$

An important special case of these hypotheses to test the significance of regression is

$$H_0 : \beta_1 = 0, \quad H_1 : \beta_1 \neq 0$$

[22, pg.24]

4.3.1.1 t-Statistic

The t-statistic defines the statistical significance of the beta coefficients. Since the estimator $\hat{\beta}_1$ is normally distributed with mean $\beta_1$ and variance $\sigma^2 / S_{xx}$, and $\sigma^2$ can be approximated with $MS_{Res}$, the t-statistic is defined as

$$t_0 = \frac{\hat{\beta}_1 - \beta_{10}}{\sqrt{MS_{Res}/S_{xx}}} = \frac{\hat{\beta}_1 - \beta_{10}}{se(\hat{\beta}_1)}$$

where $se(\hat{\beta}_1)$ is the standard error of the slope. $t_0$ follows a $t_{n-2}$ distribution if the null hypothesis defined above is true. Therefore, $t_0$ is used to test the null hypothesis by comparing the observed value of $t_0$ with the upper $\alpha/2$ percentage point of the $t_{n-2}$ distribution. The null hypothesis is thus rejected if

$$|t_0| > t_{\alpha/2,n-2}$$

If the null hypothesis $H_0 : \beta_1 = 0$ cannot be rejected, it implies that there is no linear relationship between $x$ and $y$ [22, pg.23-24].

4.3.1.2 F-Statistic

The F-statistic is an effective measure to test the significance of regression. This statistic is defined as

$$F_0 = \frac{SS_R/df_R}{SS_{Res}/df_{Res}} = \frac{SS_R}{SS_{Res}/(n-2)} = \frac{MS_R}{MS_{Res}}$$

and follows the $F_{1,n-2}$ distribution. The null hypothesis $H_0 : \beta_1 = 0$ is rejected if

$$F_0 > F_{\alpha,1,n-2}$$

[22, pg.27]
4.3.2 Confidence Intervals

The width of the confidence intervals of the parameters $\beta_0$ and $\beta_1$ is a quality measure of the model. The assumption that the errors are normally and independently distributed is again important. The sampling distribution of $(\hat{\beta}_1 - \beta_1)/se(\hat{\beta}_1)$ and $(\hat{\beta}_0 - \beta_0)/se(\hat{\beta}_0)$ is then $t$ with $n - 2$ degrees of freedom. A $100(1 - \alpha)$ percent confidence interval on the slope $\beta_1$ is thus given by

$$\hat{\beta}_1 - t_{\alpha/2, n-2}se(\hat{\beta}_1) \leq \hat{\beta}_1 \leq \hat{\beta}_1 + t_{\alpha/2, n-2}se(\hat{\beta}_1)$$

and the same interval on the intercept $\beta_0$ is

$$\hat{\beta}_0 - t_{\alpha/2, n-2}se(\hat{\beta}_0) \leq \hat{\beta}_0 \leq \hat{\beta}_0 + t_{\alpha/2, n-2}se(\hat{\beta}_0)$$

[22, pg.29]

4.4 Model Development

In order to analyze how appropriate a model is as well as develop the model’s fit, a number of strategies can be considered. Model building is an iterative approach where several measures need to be analyzed in order to improve the model.

4.4.1 Coefficient of Determination

The coefficient of determination is defined as

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_{Res}}{SS_T}$$

where

$$SS_{Res} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

and

$$SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$R^2$ is a measure of the proportion of variation explained by the regressor $x$ since $SS_T$ is a measure of the variability in $y$ without considering $x$ and $SS_{Res}$ measures the variability in $y$ remaining after $x$ has been considered. The coefficient of determination normally takes values between 0 and 1, but in rare cases it can take negative values as well. Values close to 1 imply that most of the variability in $y$ is explained by the model, which is why it is desirable for a regression model to have a large $R^2$ value [22, pg.35-36].

Similarly, $R^2_{adj}$ also gives an indication of how well the given data fits a line. The difference between the two is that $R^2_{adj}$ adjusts for the number of variables in the model.
4.4.2 Residual Mean Square

The residual mean square is also a measure used to check how well the model fits the data set. This measure represents the average variance of the data points around the fitted regression line, i.e. it is the standard deviation of residual errors. \( MS_{Res} \) is an unbiased estimator of \( \sigma^2 \), which is given by

\[
\hat{\sigma}^2 = \frac{SS_{Res}}{n - 2} = MS_{Res}
\]

The residual sum of squares has \( n - 2 \) degrees of freedom as two degrees of freedom are lost when estimating \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \). This gives us the formula for \( MS_{Res} \) presented above [22, pg.21].

A model with a small value for \( MS_{Res} \) indicates that the regression model is a good fit for the data.

4.4.3 Transformations

If the residual analysis or any of the above-mentioned measures indicate that the regression model is not a good fit, a transformation of the model may be necessary. There are multiple transformation methods depending on the fit. If the variance is not consistent, a common approach is to use variance-stabilizing transformations. A second approach is transformations to linearize the model, which is performed in the case where the model is not linear [22, pg.171].

A common transformation is the Box-Cox method, which transforms the response variable \( y \) in order to correct non-normality and/or non-constant variance. This method implies that \( y \) undergoes a power transformation, \( y^\lambda \). Using the method of maximum likelihood, \( \lambda \) can be estimated to the value for which the residual sum of squares \( SS_{Res}(\lambda) \) is a minimum. The value obtained for \( \lambda \) is then the value to be used when conducting the power transformation \( y^\lambda \) [22, pg.182-183].

5 Data

5.1 Delimitations

In order to be able to complete this study, a number of limitations were necessary to take due to the large number of funds available. Firstly, this report focuses on open-ended mutual equity funds that are actively managed, thereby excluding all sorts of index funds as well as exchange traded funds (ETFs). A second requirement for the included funds is that their inception date, the date on which the fund began its operations, must be older than three years. Lastly, each fund must have a sustainability score.

Due to the fact that there is no international standardized way to calculate the sustainability score of a fund, all funds that meet the three above-mentioned conditions were divided into four sub-groups that were individually analyzed. These geographical sub-groups are USA, Europe, the Nordic region and Asia ex-Japan. Japan is excluded from Asia since it is a highly developed economy whereas the rest of Asia is considered to be emerging markets [8].
Furthermore, a company can be categorized based on its size and its maturity. This paper includes funds that invest in small-, mid-, and large-cap as well as both value and growth companies. Hence, there are no delimitations regarding which kind of company a fund invests in.

5.2 Data Collection

5.2.1 Data Source

Morningstar Inc is a global financial services firm that provides a wide spectrum of investment research and investment management services. Their software platform, Morningstar Direct, is used in order to collect the necessary data. Its main purpose is to help professional investment managers construct new products and portfolios. Their search tool provides customized search-filters which enables tailor-made lists of funds and all relevant information about the funds.

5.2.2 Sample Selection

In order to fulfill the requirements previously specified, six different screening criteria were used when collecting the data.

The first screening criterion is to exclude all funds not listed as equity funds. In order to be categorized as an equity fund, the fund must invest at least 80% percent of its capital in equities [17]. Secondly, each fund needs to have a sustainability score. This mainly affects funds investing in small-cap companies since they are not as widely evaluated as mid- and large-cap firms when it comes to sustainability.

Including index funds might lead to the result being biased towards the ability of a fund to follow an index instead of focusing on the risk-adjusted return and its correlation with the sustainability score. Therefore, the third screening criterion is to exclude all types of funds trying to replicate an index.

The fourth screening criterion is to include funds with an inception date older than three years that are still active. This implies that funds that were merged or terminated during this time period were excluded. However, this might result in the data set to suffer from survivorship bias [27] which will be discussed in the end of this section. This criterion was necessary to include in order to explore how the funds performed during both positive and negative years, as shown by Table 1 below. The MSCI ACWI Indexes captures all sources of equity returns in 23 developed and 24 emerging markets (ACWI, 2019). The returns for the past five years are displayed in the table below [25].

<table>
<thead>
<tr>
<th>Year</th>
<th>MSCI ACWI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>-8.98%</td>
</tr>
<tr>
<td>2017</td>
<td>24.62%</td>
</tr>
<tr>
<td>2016</td>
<td>8.48%</td>
</tr>
<tr>
<td>2015</td>
<td>-1.84%</td>
</tr>
<tr>
<td>2014</td>
<td>4.71%</td>
</tr>
</tbody>
</table>

Table 1: Annual Performance of MSCI ACWI
Hence, to match the inception date mentioned above, this study includes the risk-adjusted return calculated for a three-year period.

The fifth screening criterion is that the fund must only invest in one of four earlier-specified geographical areas. This divided the funds that had passed the four earlier criteria into four subgroups, which will be analyzed individually. However, each subgroup contained several funds with the same FundID, which is a characteristic used by Morningstar to identify funds. This is because a fund might be available in different share classes targeting different groups of investors. Each share class has a specific SecID, which is a characteristic used by Morningstar to identify different share classes. The share class with the highest Total Net Assets was chosen in order to avoid including the same funds that were invested in the exact same portfolios multiple times, which is the sixth and final screening criterion.

Morningstar’s database provides information about over 80,000 open-end mutual funds. However, after implementing each respective screening criteria, the number of funds remaining in each geographical area are presented below in Table 2.

<table>
<thead>
<tr>
<th>Geographical Area</th>
<th>Number of Funds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nordic</td>
<td>103</td>
</tr>
<tr>
<td>Europe</td>
<td>575</td>
</tr>
<tr>
<td>USA</td>
<td>592</td>
</tr>
<tr>
<td>Asia</td>
<td>147</td>
</tr>
</tbody>
</table>

Table 2: Number of Funds in Each Geographical Area

5.3 Variables

5.3.1 Morningstar Portfolio Sustainability Score

The Morningstar Portfolio Sustainability Score, which is the regressor in the regression model, is an asset-weighted average of normalized company-level ESG with deductions made for controversial incidents by the issuing companies. Controversial incidents refers to discriminatory behavior, fraud and environmental accidents [24]. The Morningstar Portfolio Sustainability Score is calculated in the following way:

\[
\text{Portfolio Sustainability Score} = \text{Portfolio ESG Score} - \text{Portfolio Controversy Deduction}
\]

To make the ESG scores comparable across peer groups, which is necessary for evaluation of diversified portfolios, Morningstar normalizes the scores of each peer group. The normalized ESG scores range from 0 to 100 with a mean of 50. Once the companies’ ESG scores are normalized, they are aggregated to a portfolio ESG score using an asset-weighted average of all covered securities, which in this report only includes equity. However, to receive a Portfolio ESG Score, at least 67% of a portfolio’s assets under management must have a company ESG score [24].

Moreover, Morningstar imports data from Sustainalytics who tracks and categorizes ESG-related incidents on more than 10,000 companies globally. Because the presence of controversy is a negative contributor to a company’s overall sustainability performance, it is deducted from the overall Portfolio Sustainability Score. While a company may be involved in multiple ESG-related incidents at any given time, a company’s most serious current controversy is used to create its controversy score, on a scale from 0 to 20 [24].
In order to compute the Portfolio Sustainability Score, a number of steps are necessary. Firstly, the scores
of each peer group must be normalized using a z-score transformation. This normalization is defined
according to:

\[ Z_{\text{peer}} = \frac{ESG_x - \mu_{\text{peer}}}{\sigma_{\text{peer}}} \]

where

- \( ESG_x \) = Sustainalytics company ESG score
- \( \mu_{\text{peer}} \) = Peer group mean ESG score
- \( \sigma_{\text{peer}} \) = Peer group standard deviation of ESG scores

Secondly, the z-scores are then used to create the normalized ESG scores on a 0-100 scale, with a mean of
50, as follows:

\[ ESG_{\text{Normalized}} = 50 + (Z_{\text{peer}} \times 10) \]

The Portfolio ESG Score is then defined as:

\[ Portfolio\ ESG = \sum_{i=1}^{n} ESG_{\text{Normalized}} \times \text{Weightsadj} \]

Lastly, the Portfolio Controversy Score is an asset-weighted average of company controversy scores:

\[ MContr_p = \sum_{i=1}^{n} w_i Cont_i \]

where:

- \( MContr_p \) = the Morningstar Portfolio Controversy Score
- \( Cont_i \) = the controversy score of company \( i \)

Finally, this gives the definition of Morningstar’s Portfolio Sustainability Score:

\[ Portfolio\ Sustainability\ Score = Portfolio\ ESG\ Score - Portfolio\ Controversy\ Deduction \]

Once the Portfolio Sustainability Score has been computed, it is converted to the Morningstar Sustainability Rating. This rating is a scale from 1 to 5 and is more intuitive and easy to comprehend when comparing funds. However, since this is a simplified version of the Portfolio Sustainability Score it is not used in this report. Using this score instead would not have generated as detailed results. Nevertheless, the scale is presented below in Table 3 to give an overview of how funds are categorized in terms of sustainability.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Score</th>
<th>Descriptive Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest 10%</td>
<td>5</td>
<td>High</td>
</tr>
<tr>
<td>Next 22.5%</td>
<td>4</td>
<td>Above Average</td>
</tr>
<tr>
<td>Next 35%</td>
<td>3</td>
<td>Average</td>
</tr>
<tr>
<td>Next 22.5%</td>
<td>2</td>
<td>Below Average</td>
</tr>
<tr>
<td>Lowest 10%</td>
<td>1</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 3: Morningstar Sustainability Rating

5.3.2 Risk-Adjusted Return

Throughout this thesis, the Sharpe ratio is used as a measurement of risk-adjusted return and it is the response variable in the regression model. Morningstar calculates the Sharpe ratio in the following way:
First, the average monthly return of the 90-day Treasury bill (over a 36-month period) is subtracted from the excess return of a fund beyond that of the 90-day Treasury bill, a risk-free investment. An arithmetic annualized excess return is then calculated by multiplying this monthly return by 12. To show a relationship between excess return and risk, this number is then divided by the standard deviation of the annualized excess returns of the fund [23].

5.4 Critics of the Data Set

As previously described under section 5.2.2, all funds that have been terminated or merged during the specified time period were excluded from the data. Thereby, all funds that performed extremely poorly during the time period were not represented in the data. Consequently, the results may suffer from survivorship bias [27]. Nevertheless, this was necessary since this paper is dependent on data from Morningstar and "dead funds" without a Sharpe ratio and Portfolio Sustainability Score would not have provided any useful data.

Furthermore, the Portfolio Sustainability Score that funds receive from Morningstar is a static score and does not reflect any previous years. This is because the underlying holdings of a fund are exchanged over time and since the score is based on the underlying holdings, it has to be re-evaluated when a fund reconstructs its portfolio. However, it is shown that the persistence of ESG scores in mutual funds is approximately two years and that the score is terminated after three [29]. Thereby, the Portfolio Sustainability Score for each fund is representable for the period which the Sharpe ratio is calculated for. This way, the results should be of significance, and the sustainability ratings to be relevant. In conclusion, this paper used the Morningstar Portfolio Sustainability Score for all funds as of March 2019, and assumes this to be the correct rating for the entire three-year period.

Finally, the funds included in each respective geographical sub-category are not traded in the same currency. Since currencies are constantly exposed to fluctuations in exchange rates, they become inherently volatile, making holders of a given currency vulnerable to its depreciation against other currencies. Thereby, the risk-adjusted return might be affected by currency exposure. However, an assumption has been made that the effect on the risk-adjusted return is equally distributed among the different portfolio sustainability scores. Hence, it should not significantly affect the results.

6 Analysis

6.1 Nordic Region

6.1.1 Residual Analysis

- Normal Probability Plot and Residual vs. Fitted Values Plot

Figure 2(a) below demonstrates that the data set for the Nordic Region is approximately normally distributed, as the data points generally follow a straight line. The three observations that deviate the most from the normal fit are observations 13, 48 and 58. As the red line in figure 2(b) is not exactly horizontal, it is clear that there is not equal variance along the regression line. However, the data is more or less symmetric around the origin. The same three data points that deviate from the normal distribution in figure 2(a) also differ from the trend in 2(b).
Figure 2: Residual Plots for the Nordic Region

• Residuals

Figure 3 includes the standardized, studentized, PRESS and R-student residuals for the Nordic Region. As stated earlier under section 4.2.1, if the data set is large enough, the residual plots generally provide the same information. However, since a "large data set" is a vague expression and is not clearly defined, all four residual plots were examined and compared in order to obtain a detailed analysis. This has been done in each analysis in the respective regions.

In figure 3(a), representing the standardized residuals, it is clear that there is no observation that has a standardized residual value greater than the absolute value of three. This indicates that there are no points that are considered to be outliers. Similarly, the plot of the studentized and R-student residuals follow the same trend and shape as the standardized residuals.

Figure 3(c), representing the PRESS residuals, shows that there are three observations that have larger PRESS residuals than the rest, indicating that they are more influential. These observations are points 13, 48 and 58.

Figure 3: Scaled Residuals for the Nordic Region
• **Adjusted-Variable Plot**

Since the Adjusted-Variable Plot is not a horizontal line, there is a linear relationship between the regressor and response variable. Specifically, there is a negative relationship between the two.

![Adjusted-Variable Plot](image)

**Figure 4: Adjusted-Variable Plot for the Nordic Region**

**6.1.2 Cook’s Distance**

In the plot below there are three points that have a larger Cook’s Distance than the rest, including observations 11, 13 and 58. However, since the distance is remarkably smaller than the limit of 1, none of the observations are considered to be highly influential.

![Cook's Distance Plot](image)

**Figure 5: Cook’s Distance Plot for the Nordic Region**

**6.1.3 CovRatio**

The CovRatio Plot below shows the lower cutoff line, marked in red. There are two points that are considered to have considerable influence on the Covariance Matrix, including number 13 and 58.
6.1.4 Model Properties

- **t-Test**

As the intercept has a larger t-statistic in absolute terms than the slope, it indicates that the intercept is more significant than the slope. However, since both variables are highly significant, the null hypothesis can be rejected, indicating that there is a significant association between the regressor and response variable.

<table>
<thead>
<tr>
<th></th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>7.729</td>
</tr>
<tr>
<td>Slope</td>
<td>-4.217</td>
</tr>
</tbody>
</table>

Table 4: t-Statistic for the Nordic Region

- **F-Statistic**

A large F-Statistic indicates a statistically significant regression model. The regression model in the Nordic region obtained a F-statistic equal to 17.78, indicating that the model is significant.

- **Adjusted $R^2$**

The regression model has an adjusted $R^2$ value of 0.1413. Since it is desirable to have a value close to 1, this metric indicates that a large proportion of the variability in the outcome can not be explained by the regression model. This implies that regression model does not fit the data set particularly well.

- **Residual Mean Square**

The model has a residual mean square that is equal to 0.1963. Since this value is close to zero, it is an indication that the residuals have a low standard deviation.

- **Confidence Interval**

The table below presents the 95% confidence interval for both the intercept and slope in the regression model. Since the limits of the confidence interval for the slope are both negative, it gives an indication that the slope most certainly is negative.
<table>
<thead>
<tr>
<th></th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.69296192</td>
<td>2.86205965</td>
</tr>
<tr>
<td>Slope</td>
<td>-0.03349481</td>
<td>-0.01206341</td>
</tr>
</tbody>
</table>

Table 5: Confidence Intervals for the Nordic Region

6.1.5 Box Cox

In the plot below, the maximum value of $\lambda$ is roughly $\lambda = 1.35$, which is the most optimal value of lambda and could be used in a power transformation. However, the plot also shows a 95 percent confidence interval for $\lambda$, and it is obvious that $\lambda = 1$ is included in this interval. This, in combination with the residual analysis and model properties obtained above, indicates no transformation of the regression model is necessary.

Figure 7: Box Cox Transformation Plot for the Nordic Region

6.2 Europe

6.2.1 Residual Analysis

- Normal Probability Plot and Residual vs. Fitted Values Plot

Since the data points in figure 8(a) generally follow a straight line, it is clear that the European data set is normally distributed. The three observations that deviate the most from the normal fit are observations 11, 40 and 357. Next, since the red line in figure 8(b) below is approximately horizontal, it indicates that there is more or less equal variance along the regression line. Furthermore, the data is symmetrically spread around the origin. There are, however, a few points that do not follow the same pattern as most, and this includes the exact same points mentioned above that deviate from the normal distribution.

Figure 8: Residual Plots for Europe
• Residuals
As a rule of thumb, the value of the standardized residuals should not exceed the absolute value of three. In figure 9 (a), it is clear that there are a number of points that are close to this limit, and three points exceeding it. These points include the same three points as above, namely 11, 40 and 357. Therefore, these three points may be considered to be outliers. Similarly, the plots of the studentized and R-student residuals follow the same shape as the standardized residuals. Figure 9(c), representing the PRESS residuals, demonstrates that there are three observations that have larger PRESS residuals than the rest, indicating that they are more influential. These observations are the same points as above.

![Graphs of residuals](image)

(a) Standardized Residuals  (b) Studentized Residuals  
(c) PRESS Residuals  (d) R-Student Residuals

Figure 9: Scaled Residuals for Europe

• Adjusted-Variable Plot
If the Adjusted-Variable Plot would exactly be a horizontal line, it would indicate that there is no linear relationship between the regressor and response variable. In the plot below, the line is almost horizontal, but it is slightly negative. Therefore, there is a minor negative relationship between the two variables.

![Adjusted-Variable Plot](image)

Figure 10: Adjusted-Variable Plot for Europe
6.2.2 Cook’s Distance

In the plot below there are three points that have a larger Cook’s Distance than the rest, including observations 311, 357 and 486. However, since the distance is remarkably smaller than the limit of 1, none of the observations are considered to be outliers.

![Cook's Distance Plot for Europe](image)

Figure 11: Cook’s Distance Plot for Europe

6.2.3 CovRatio

The CovRatio Plot below shows the upper and lower cutoff line, marked in red. There are 49 points, marked in red, that are outside of these cutoff lines, indicating that these points have a greater influence on the Covariance Matrix than the rest. However, since the previous methods used indicate that there are no outliers in the model, this result only implies that these 49 points may be more influential than the rest, but it does not mean that the points greatly deviates from the rest of the data.

![CovRatio Plot for Europe](image)

Figure 12: CovRatio Plot for Europe

6.2.4 Model Properties

- t-Test

As the intercept has a larger t-statistic in absolute terms than the slope, it indicates that the intercept is more significant than the slope. Since the slope’s t-statistic is quite low, it is not considered to be highly significant, and the null hypothesis can not be rejected. If we cannot reject the null hypothesis, it implies
that we cannot confidently state that there is a significant association between the regressor and response variable.

<table>
<thead>
<tr>
<th></th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>4.307</td>
</tr>
<tr>
<td>Slope</td>
<td>-0.601</td>
</tr>
</tbody>
</table>

Table 6: t-Statistic for Europe

- **F-Statistic**

A large F-Statistic indicates a statistically significant regression model. The European regression model obtained a F-statistic equal to 0.3616, indicating that the model is not significant.

- **Adjusted $R^2$**

The regression model has an adjusted $R^2$ value of -0.001113. It is desirable to have a value close to 1, as this would indicate that a large proportion of the variability in the outcome can be explained by the regression model. A negative adjusted $R^2$ value of -0.001113 implies that the regression model does not reflect the data set and that the correlation between the regressor and response variable is negligible.

- **Residual Mean Square**

The model has a residual mean square that is equal to 0.1802. Since this value is close to zero, it is an indication that the residuals have a low standard deviation.

- **Confidence Interval**

The table below presents the 95% confidence interval for both the intercept and slope in the regression model. Since the lower limit of the confidence interval for the slope is negative whilst the upper limit is positive, it supports that there is no clear relationship between the regressor and response variable.

<table>
<thead>
<tr>
<th></th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.336898331</td>
<td>0.901810282</td>
</tr>
<tr>
<td>Slope</td>
<td>-0.006788118</td>
<td>0.003605763</td>
</tr>
</tbody>
</table>

Table 7: Confidence Intervals for Europe

6.2.5 **Box Cox**

In the plot below, the maximum value of $\lambda$ is roughly $\lambda = 0.8$, which is the most optimal value of lambda and could be used in a power transformation. However, the plot also shows a 95 percent confidence interval for $\lambda$, and it is obvious that $\lambda = 1$ is included in this interval. This, in combination with the residual analysis and model properties obtained above, indicates no transformation of the regression model is necessary.
6.3 USA

6.3.1 Residual Analysis

- Normal Probability Plot and Residual vs. Fitted Values Plot

Figure 14(a) shows that the American data set generally follows a straight line, indicating that the data is normally distributed. There are three observations that slightly deviate from the normal fit, and these observations are 199, 235, and 330. Since the red line in plot 14(b) is almost exactly horizontal, it indicates that there is equal variance along the regression line. Furthermore, the data is symmetrically spread around the origin. There are, however, a few points that do not follow the same pattern as most. These points are the same as the points that deviate from the normal fit observed earlier.

- Residuals

The value of the standardized residuals should not exceed the absolute value of three. There are a few points, as shown in figure 15(a) that are close to or exceeds this limit. Therefore, such points are considered to be more influential. Similarly, figure 15(b) and figure 15(d) representing the studentized and R-student residuals respectively, show the same results.

Furthermore, figure 15(c) demonstrates that there are a few observations that have larger PRESS residuals than the rest, indicating that they are more influential. However, these residuals still have low residual values, and thus they are not unusually influential.
Adjusted-Variable Plot

Since the line in the Adjusted-Variable Plot has a positive slope, it indicates that there is a positive relationship between the regressor and response variable. The plot below indicates a quite clear correlation between the two.

Figure 16: Adjusted-Variable Plot for USA

6.3.2 Cook’s Distance

In the plot below there are three points that have a larger Cook’s Distance than the rest, including observations 235, 330 and 383. However, since the distance is remarkably smaller than the limit of 1, none of the observations are considered to be outliers.
6.3.3 CovRatio

The CovRatio Plot below shows the upper and lower cutoff line, marked in red. There are 26 points, marked in red, that are outside of these cutoff lines, indicating that these points have a greater influence on the Covariance Matrix than the rest. However, since the previous methods used indicate that there are no outliers in the model, this result only implies that these 26 points may be more influential than the rest, but it does not mean that the points greatly deviates from the rest of the data.

![CovRatio Plot for USA](image1)

Figure 18: CovRatio Plot for USA

6.3.4 Model Properties

- t-Test

Since the slope has a larger t-statistic in absolute terms than the intercept, it indicates that the slope is more significant than the intercept. Since the slope’s t-statistic is relatively high, it is considered to be highly significant. Therefore, the null hypothesis can be rejected, indicating that there is a significant association between the regressor and response variable.

<table>
<thead>
<tr>
<th></th>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.795</td>
</tr>
<tr>
<td>Slope</td>
<td>9.607</td>
</tr>
</tbody>
</table>

Table 8: t-Statistic for USA
• **F-Statistic**

A large F-Statistic indicates a statistically significant regression model. The American regression model obtained a F-statistic equal to 92.29, which is quite large, indicating that the model is significant.

• **Adjusted $R^2$**

The regression model has an adjusted $R^2$ value of 0.1338. Since it is desirable to have a value close to 1, this metric indicates that a large proportion of the variability in the outcome cannot be explained by the regression model. This implies that the regression model does not fit the data set very well.

• **Residual Mean Square**

The model has a residual mean square that is equal to 0.206. Since this value is relatively close to zero, it is an indication that the residuals have quite a low standard deviation.

• **Confidence Interval**

The table below presents the 95% confidence interval for both the intercept and slope in the regression model. Since the limits of the confidence interval for the slope are both positive, it gives an indication that the slope most certainly is positive.

<table>
<thead>
<tr>
<th></th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.47853895</td>
<td>0.02146413</td>
</tr>
<tr>
<td>Slope</td>
<td>0.02204908</td>
<td>0.03338097</td>
</tr>
</tbody>
</table>

Table 9: Confidence Intervals for USA

6.3.5 **Box Cox**

In the plot below, the maximum value of $\lambda$ is exactly $\lambda = 1$, which directly indicates that no transformation is needed. The plot also shows a 95 percent confidence interval for $\lambda$, and this interval is roughly from 0.75 to 1.25. Due to the obtained plot, it is obvious that no transformation is necessary.
6.4 Asia ex-Japan

6.4.1 Residual Analysis

- Normal Probability Plot

Even though the plot in figure 20(a) is not at straight as the earlier normal probability plots, the data generally follows a straight line, indicating that the Asian data is also normally distributed. The three observations that deviate the most from the normal fit are observations 68, 121 and 135. The red line in figure 20(b) is not horizontal, which indicates that there is not equal variance along the regression line. However, the data is generally symmetrically spread around the origin.

![Normal Probability Plot](image1)

(a) Normal Probability Plot

![Residual vs. Fitted Values Plot](image2)

(b) Residual vs. Fitted Values Plot

Figure 20: Residual Plots for Asia ex-Japan

- Residuals

In figure 21(a) it is clear that there are no points that exceed the absolute value of three. Therefore, there are no points in the given data set that are considered to be outliers. The studentized and R-student residuals show similar plots, indicating that the same conclusion can be made. Figure 21(c) demonstrates a large spread of the PRESS residuals in comparison to the earlier geographical regions. There are a number of points that have larger PRESS residuals than the rest, indicating that they are more influential. These residuals still have low residual values, however, indicating that they are not noticeably influential.

![Standardized Residuals](image3)

(a) Standardized Residuals

![Studentized Residuals](image4)

(b) Studentized Residuals

![PRESS Residuals](image5)

(c) PRESS Residuals

![R-Student](image6)

(d) R-Student Residuals

Figure 21: Scaled Residuals for Asia ex-Japan
• **Adjusted-Variable Plot**

If the Adjusted-Variable Plot would exactly be a horizontal line, it would indicate that there is no linear relationship between the regressor and response variable. In the plot below, the line has a positive slope, indicating that there is a positive relationship between the two variables.

![Adjusted-Variable Plot for Asia ex-Japan](image)

**Figure 22: Adjusted-Variable Plot for Asia ex-Japan**

### 6.4.2 Cook's Distance

In the plot below there are three points that have a larger Cook's Distance than the rest, including observations 12, 30 and 68. However, since the distance is remarkably smaller than the limit of 1, none of the observations are considered to be outliers.

![Cook's Distance Plot for Asia ex-Japan](image)

**Figure 23: Cook's Distance Plot for Asia ex-Japan**

### 6.4.3 CovRatio

The CovRatio Plot below shows the upper and lower cutoff line, marked in red. There are 14 points, marked in red, that are outside of these cutoff lines, indicating that these points have a greater influence on the Covariance Matrix than the rest.

![CovRatio Plot](image)
6.4.4 Model Properties

- **t-Test**

As the slope has a larger t-statistic in absolute terms than the intercept, it indicates that the slope is more significant than the intercept. However, since the slope’s t-statistic is quite low, it is not considered to be highly significant, and the null hypothesis can not be rejected. Since we cannot reject the null hypothesis, it implies that we cannot confidently state that there is a significant association between the regressor and response variable.

<table>
<thead>
<tr>
<th>t-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>Slope</td>
</tr>
</tbody>
</table>

Table 10: t-Statistic for Asia ex-Japan

- **F-Statistic**

A large F-Statistic indicates a statistically significant regression model. The Asian regression model obtained a F-statistic equal to 4.757, indicating that the model is not particularly significant.

- **Adjusted $R^2$**

The regression model has an adjusted $R^2$ value of 0.02509. It is desirable to have a value close to 1, as this would indicate that a large proportion of the variability in the outcome can be explained by the regression model. Since the obtained value is close to zero, it implies that the regression model does not reflect the data set particularly well and that the correlation between the regressor and response variable is not very strong.

- **Residual Mean Square**

The model has a residual mean square that is equal to 0.1778. Since this value is close to zero, it is an indication that the residuals have a low standard deviation.

- **Confidence Interval**

The table below presents the 95% confidence interval for both the intercept and slope in the regression model. Since both the lower and upper limit of the confidence interval for the slope is positive, the
relationship between the regressor and response variable is most likely a positive one. However, the lower limit is very close to zero, indicating that this relationship is very low.

<table>
<thead>
<tr>
<th></th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.943323845</td>
<td>0.74089146</td>
</tr>
<tr>
<td>Slope</td>
<td>0.001961951</td>
<td>0.03985154</td>
</tr>
</tbody>
</table>

Table 11: Confidence Intervals for Asia ex-Japan

6.4.5 Box Cox

In the plot below, the maximum value of $\lambda$ is roughly $\lambda = 1.55$, which is the most optimal value of lambda and could be used in a power transformation. However, the plot also shows a 95 percent confidence interval for $\lambda$, and the lower limit is on $\lambda = 1$. This, in combination with the residual analysis and model properties obtained above, indicates that no transformation of the regression model is necessary.

![Box Cox Transformation for Asia ex-Japan](image)

Figure 25: Box Cox Transformation for Asia ex-Japan

7 Results

7.1 Final Models

Four regression models have been obtained; one for each respective geographical region. Table 12 below presents the values of the intercept and slope coefficients in each regression model.

<table>
<thead>
<tr>
<th>Region</th>
<th>Intercept</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nordic</td>
<td>2.277511</td>
<td>-0.022779</td>
</tr>
<tr>
<td>Europe</td>
<td>0.619354</td>
<td>-0.001591</td>
</tr>
<tr>
<td>USA</td>
<td>-0.228537</td>
<td>0.027715</td>
</tr>
<tr>
<td>Asia ex-Japan</td>
<td>-0.101216</td>
<td>0.027715</td>
</tr>
</tbody>
</table>

Table 12: Coefficient Values for each Region

Figure 26 represents the final regression model obtained for each geographical region. The grey area around the regression lines represents the confidence interval for each model.
Finally, each final regression model is presented below.

- **Final Model for Nordic Region**
  \[
  \text{Sharpe Ratio} = 2.277511 - 0.022779 \times \text{Sustainability Score}
  \]

- **Final Model for Europe**
  \[
  \text{Sharpe Ratio} = 0.619354 - 0.001591 \times \text{Sustainability Score}
  \]

- **Final Model for USA**
  \[
  \text{Sharpe Ratio} = -0.228537 + 0.027715 \times \text{Sustainability Score}
  \]

- **Final Model for Asia ex-Japan**
  \[
  \text{Sharpe Ratio} = -0.101216 + 0.020907 \times \text{Sustainability Score}
  \]
8 Discussion

8.1 Interpretation of Results and Previous Studies

The results indicate that within the geographical regions USA and Asia ex-Japan, there is a positive correlation between the Portfolio Sustainability Score and the risk-adjusted return. The opposite was shown for the Nordic region and Europe where there instead was a negative correlation between the two.

However, for all four of the regions, the adjusted $R^2$ values are noticeably low which implies that the regression models do not reflect the different data sets particularly well. Furthermore, the results in Europe and Asia ex-Japan were not of statistical significance, indicating that no relationship was found in these regions. Even though the results in the Nordic region and USA were of significance, the correlation was not particularly strong.

Therefore, this paper has not found any support for the idea that sustainable investments yield a higher risk-adjusted return. Nevertheless, it has shown that investing in sustainable funds does not compromise on the financial return. Therefore, the results suggest that it is neither better nor worse to invest in sustainable funds compared to conventional ones. This leads to two questions regarding the obtained results: why is there no difference with regards to return when investing in sustainable and conventional funds, and why do sustainable funds not outperform conventional funds?

After negative screening is performed, sustainable funds have a smaller investment universe than conventional funds. Intuitively, this might lead to the conclusion that sustainable funds are inferior to conventional funds, and would thus generate a lower risk-adjusted return. However, this report, as well as many others, have shown otherwise. One potential reason could be that the limited investment universe for sustainable funds is still large enough to avoid considerable loss in diversification. Accordingly, sustainable funds do not expose themselves to a higher risk and gain approximately equal risk-adjusted returns as that of conventional funds.

On the other hand, one plausible reason for why sustainable funds do not outperform conventional funds might be the lack of small-cap companies in their portfolios, since small-cap firms often do not have a sustainability rating. It has been proven that over time, small-cap firms generate greater returns than mid- and large-cap firms [26]. Thereby, if the same analysis had been done with only mid- and large-cap companies, the results might had tilted in favor of sustainable funds instead. However, it is worth mentioning that this is only a speculation and the answer is left to further studies in the area.

This paper found a positive correlation between the risk-adjusted return and the sustainability score of a mutual fund in USA and Asia ex-Japan and the opposite in the Nordic region and Europe. Even though the results were not of statistical significance, this outcome might seem odd since SRI is much more common in Europe and the Nordic region. The explanation can perhaps be found in Figure 26 if it is studied carefully. For USA and Asia ex-Japan, the sustainability scores range from below 40 to approximately 50. Meanwhile, in the Nordic region and Europe, they range from just below 50 up to 60. This indicates that a fund with a low sustainability score in Europe and the Nordic region is considered to be a very sustainable fund in USA and Asia ex-Japan. One idea could be that investing in sustainable companies yields a higher return, but only up to a certain limit. In this case, the limit is somewhere around 50. After that limit is passed, a higher sustainability score compromises the risk-adjusted return due to a more narrow investment universe. However, this is again only a speculation and there is no concrete evidence supporting this idea.

A limitation of this study is the lack of time series data on the sustainability ratings provided by Morn-
ingstar. This report was forced to focus on a relatively short time frame since Morningstar only provides the latest Portfolio Sustainability Score, which only represents the last three years. Thereby, this report only provides insight regarding development in the near future. However, if an investor intends to hold an investment for a long time, it might be relevant to research how the risk-adjusted return correlates with the sustainability score over longer time periods.

8.2 Implementation for Carnegie Fonder

8.2.1 How Carnegie Fonder Works with Sustainability Today

In order to establish how Carnegie Fonder can improve their sustainability work, an extensive analysis of how they currently work had to be carried out. This section is primarily based on information from their website as well as an interview with Erik Amcoff, Head of Communications at Carnegie Fonder.

To begin with, Carnegie Fonder’s sustainability analysis originates from ESG-related mega-trends, such as climate changes, technological paradigm changes and resource shortage. Carnegie Fonder uses these, in combination with financial data, to analyze how well a company will perform in the future and whether the investment will yield a high risk-adjusted return in the long run. Furthermore, the foundation of their sustainability work is the UN Global Compact as well as the Principles for Responsible Investments (PRI). Carnegie Fonder has an ongoing dialogue with all of their holdings and uses data from GES International to ensure that their holdings fulfill these requirements. If a company violates these international conventions, Carnegie Fonder tries to influence them to handle these violations and improve their sustainability work. However, if the company is reluctant to improve, they remove the holding from their portfolio [15].

Carnegie Fonder uses negative screening to exclude industries that are not in line with their philosophy. When asked if they use positive screening to find companies who perform extraordinarily well within ESG, Erik Amcoff answered "We invest in well-managed value companies, so when we find a company that fulfills our values they often fulfill the ESG requirements as well. Therefore, there is no need to explicitly use positive screening.” [2]

According to Erik Amcoff, the main reason for incorporating SRI at Carnegie Fonder is to manage risk and not necessarily to find value. To do this effectively, all fund managers have taken part in two courses. In 2017, the basic course RI Fundamentals was completed, which was then complemented with an in-depth course in 2018 about responsible investments, called PRI Academy’s RI Essentials [2].

The fund managers are individually responsible for ensuring that the investments they pursue are sustainable. However, they are supported by Carnegie Fonder’s sustainability council who help the fund managers integrate SRI by providing guidance, education and necessary tools [14]. The members of the sustainability council are:

- **Hans Hedström**: CEO & CIO
- **Svante Lundberg**: Compliance Officer
- **Peter Gullmert**: Head of Sales
- **Erik Amcoff**: Head of Communications
- **Karin Fries**: Fund Manager

Lastly, Carnegie Fonder reports the carbon footprint of their equity funds and thereby makes it possible for unit holders to compare them with other funds. This measurement is a tool for the fund managers and can be used in ongoing discussions with the holdings. As recommended by the Swedish Investment Fund Association, Carnegie Fonder uses a measurement that shows the carbon intensity of their equity funds. A
low rating implies that the fund invests in companies whose operations produce low emissions in relation to revenues to a significant extent. The measure indicates the carbon footprint of holdings in relation to net sales (tonnes/SEK millions), weighted according to the holdings’ percentage of the relevant fund portfolio [12]. This is Carnegie Fonder’s form of public reporting of ESG, which creates accountability and transparency.

8.2.2 How Carnegie Fonder can Improve Their Sustainability Work

This section provides guidelines on how Carnegie Fonder can improve their implementation of SRI, based on how they currently work with ESG in combination with the findings of this report as well as previous research. Specifically, research by McKinsey&Company and their report From "why" to "why not" is included.

Erik Amcoff stated, during the interview, that there has not been a significant increase in demand of sustainable investing among their private clients as a result of the sustainability trend. However, this may be reasonable since the demand of sustainable investing strategies particularly appeal young generations and the average age of Carnegie Fonder’s private clients is relatively high. Therefore, in order to maintain their status as the highest ranked fund company in Sweden [13], it is essential that they improve their integration of SRI in order to attract a broader range of clients.

When SRI was initially introduced as a concept, negative screening was the main approach. Negative screening remains a relevant strategy today as it accounts for two-thirds of sustainable investments. However, more sophisticated strategies have recently emerged where there is typically less emphasis on ethical concerns and instead a larger focus on achieving a conventional investment goal: maximizing risk-adjusted returns. When an investor integrates ESG factors into the investment process without relying on time-tested standard practices, the results can be compromised [6].

As mentioned in the previous section 8.2.1, Carnegie Fonder mainly uses negative screening and their primary focus is to strengthen risk management rather than create value. This strategy is designed to exclude companies, sectors and geographies that are considered particularly risky with regards to ESG factors. Carnegie Fonder also uses proactive engagement by maintaining a dialogue with corporate managers to ensure a high level of ESG integration. Lastly, they have integrated Principles for Responsible Investment (PRI) and the UN Sustainable Development Goals (SDG) into their business. Certain SDGs are prioritized and linked to their investment strategies in order to improve corporate performance in those specific areas. All of the above are actions that Carnegie Fonder has incorporated well into their business in order to to maximize returns without undue risk of loss, as well as contribute to their risk-management.

It has been proven that when the core activities of sustainable investing are integrated into existing processes, instead of carried out in parallel, the strategy becomes more effective. To succeed in this, sustainable investment strategies must derive from an institution’s overall mandate where specific ESG objectives should be considered [6]. Carnegie Fonder’s investment mandate is Long-term, focused value investing. Carnegie Fonder defines value investing in the following way: "Value investing means that we prefer well-managed companies with established operations, sustainable business models, strong financials, a great track record and a high dividend yield. Quite simply, value companies." [11] Thereby, sustainable investments are an underlying part of their overall mandate.

Furthermore, defining and tracking progress against clear metrics and targets is vital in order to be able to measure the success of the sustainable investment strategy. For example, this might include measuring the proportion of the portfolio managed with respect to ESG factors, measuring the ratios between executive pay and worker pay, or measuring carbon emission reductions [6]. Currently, Carnegie Fonder only does the latter, and this is something that could be further developed to increase transparency in their public
reporting. This would also contribute to their performance management since it would facilitate the follow-up of the performance of internal managers with regards to these aspects.

As a fund company, it is important to identify what expertise is needed to carry out the sustainable investment strategy. For instance, those that actively engage with management teams may need specialists with executive experience. Companies that rely on screening techniques benefit from expertise in quantitative analysis [6]. Carnegie Fonder uses an in-house ESG team, as previously mentioned. In addition, they use commercial databases to obtain high-quality information regarding companies' ESG performance.

It is also utterly important for a fund company to report on their sustainable investing practices and performance. There are several different levels of how elaborate this report should and must be. The goal of reporting on ESG performance determines which level a specific fund company finds itself in. As mentioned previously in this section, younger generations is the main age group who feel that ESG is an important aspect to take into account when making investment decisions. Today, Carnegie Fonder provides information on the carbon footprint of each equity fund. In addition, each respective fund manager discloses in their annual reports how they have incorporated sustainability and how it has affected their fund. The reports also describe what actions they have taken in general, including for example which sectors have been excluded when using negative screening. However, in order to be able to attract future private clients who demand more when it comes to sustainability, Carnegie Fonder must develop their current routines. One idea that was presented during the interview with Erik Amcoff is to introduce ESG scorecards for each fund. This would give a detailed overview of each respective fund and indicate how well they perform in terms of ESG. This would enable Carnegie Fonder to present their rigorous work within sustainability in more transparent way and would also allow private investors to systematically compare them with competitors.

As this report concluded, as well as many others, there does not seem to be a correlation between the risk-adjusted return of a fund and its overall sustainability score. Thereby, by focusing on using SRI for risk management rather than value creation seems to be the best strategy for Carnegie Fonder.

To summarize, Carnegie Fonder currently incorporates SRI well into their investing practices and their business model overall. Their main focus should onward be to improve their transparency regarding how their funds perform in terms of ESG, and not only rely on third-party providers of these metrics. As mentioned earlier, this could be done by developing ESG scorecards for each fund. As a result, this could attract younger generations who do not choose funds solely based on financial performance, but also their performance in sustainability.

8.3 Conclusions

As stated in the introduction of this thesis, the aim was to determine whether sustainable investing enhances the risk-adjusted return. This paper has examined conventional and sustainable open-ended mutual equity funds in four different geographical regions in order to reach clarification on this topic.

To begin with, the findings in each geographical region slightly differ. In Asia ex-Japan and USA, the correlation between the risk-adjusted return of a mutual fund and its sustainability score was found to be positive. On the contrary, the relationship is negative in Europe and the Nordic region. However, it is important to note that the obtained results were not of statistical significance in two of the regions, meaning that there is no relationship between the financial performance of a fund and its sustainability. In the other two regions, the correlation between the two variables was not particularly strong. Nevertheless, this study is an argument against excluding sustainable investments due to anticipated lower financial returns. The obtained results are of interest to anyone looking to invest in open-ended mutual equity funds, as they reveal that there is no additional cost of investing in a sustainable manner versus a conventional one. Since
the risk-adjusted returns between sustainable and conventional funds are equivalent, this advocates for that both conscious as well as indifferent investors should prefer to invest in a sustainable manner.

This thesis was carried out in collaboration with the fund management firm Carnegie Fonder. Carnegie Fonder already today incorporates SRI well into their business model - their main purpose being to manage risk rather than enhance value. This approach, given the results of this study as well as earlier studies, seems to be a favorable one. However, since the demand of sustainable investing has rapidly grown the past decades, especially among younger generations, Carnegie Fonder is encouraged to increase their transparency regarding the performance of a fund with regards to sustainability.
References


Appendices

A List of Tables and Figures

List of Tables

1 Annual Performance of MSCI ACWI ........................................... 20
2 Number of Funds in Each Geographical Area ................................. 21
3 Morningstar Sustainability Rating .............................................. 22
4 t-Statistic for the Nordic Region ............................................ 26
5 Confidence Intervals for the Nordic Region ................................. 27
6 t-Statistic for Europe ......................................................... 30
7 Confidence Intervals for Europe ............................................. 30
8 t-Statistic for USA .......................................................... 33
9 Confidence Intervals for USA ............................................... 34
10 t-Statistic for Asia ex-Japan ................................................. 37
11 Confidence Intervals for Asia ex-Japan ................................... 38
12 Coefficient Values for each Region ........................................ 38

List of Figures

1 Patterns for Residual vs. Fitted Values Plots ............................... 14
2 Residual Plots for the Nordic Region ...................................... 24
3 Scaled Residuals for the Nordic Region .................................. 24
4 Adjusted-Variable Plot for the Nordic Region ............................ 25
5 Cook’s Distance Plot for the Nordic Region ............................... 25
6 CovRatio Plot for the Nordic Region ...................................... 26
7 Box Cox Transformation Plot for the Nordic Region ..................... 27
8 Residual Plots for Europe .................................................... 27
9 Scaled Residuals for Europe ............................................... 28
10 Adjusted-Variable Plot for Europe ......................................... 28
11 Cook’s Distance Plot for Europe .......................................... 29
12 CovRatio Plot for Europe .................................................. 29
13 Box Cox Transformation for Europe ....................................... 31
14 Residual Plots for USA ..................................................... 31
15 Scaled Residuals for USA .................................................. 32
16 Adjusted-Variable Plot for USA ........................................... 32
17 Cook’s Distance Plot for USA .............................................. 33
18 CovRatio Plot for USA ..................................................... 33
19 Box Cox Transformation Plot for USA .................................... 34
20 Residual Plots for Asia ex-Japan ........................................... 35
21 Scaled Residuals for Asia ex-Japan ....................................... 35
22 Adjusted-Variable Plot for Asia ex-Japan ................................ 36
23 Cook’s Distance Plot for Asia ex-Japan ................................ 36
24 CovRatio Plot for Asia ex-Japan ......................................... 37
25 Box Cox Transformation for Asia ex-Japan .............................. 38
26 Final Models ................................................................. 39
B  Source Code in R

library(MASS)
library(car)
library(glmnet)
library(stats4)
library(ggplot2)

#Import data for specific geographical region
asia = read.csv("Asia.csv")

x <- as.matrix(asia[,c("Sustainability.Score")])
y <- as.matrix(asia[,c("Sharpe.Ratio")])

fit <- lm(y ~ x, data=asia)

plot(fit, 1:6)
summary(fit)
confint(fit)
ggplot(asia, aes(x = Sustainability.Score, y = Sharpe.Ratio)) +
  geom_point() +
  stat_smooth(method = "lm", col = "red")

#Normal Residual
res <- resid(fit)
plot(res,
     ylab="Residuals",
     xlab="Observation Number",
     main="Normal Residuals")
abline(0, 0)

#Standardized Residual
stdres <- rstandard(fit)
plot(stdres,
     ylab="Residuals",
     xlab="Observation Number",
     main="Standardized Residuals")
abline(0, 0)

#Studentized Residual
stud <- studres(fit)
plot(stud,
     ylab="Residuals",
     xlab="Observation Number",
     main="Studentized Residuals")
abline(0, 0)

#PRESS
r <- resid(fit)
pr <- r/(1 - lm.influence(fit)$hat)
press <- pr^2
sumpress <- sum(press)
plot(press,
     ylab="Residuals",
     xlab="Observation Number",
     main="PRESS Residuals")

#R-STUDENT
rstud <- rstudent(fit)
plot(rstud,
     ylab="Residuals",
     xlab="Observation Number",
     main="R-Student")
abline(0, 0)

#Adjusted–variable plot
avPlots(fit)

#Cook’s distance
cookd <- cooks.distance(fit)
plot(cookd,
     ylab="Cook’s Distance",
     xlab="Observation Number",
     main="Cook’s Distance")

#Covariance Ratio
covr <- covratio(fit)
p = ncol(asia)
n = nrow(asia)
cutoff_pos = 1 + 3*(p/n)
cutoff_neg = 1 - 3*(p/n)
plot(covr, pch="*", cex =2,
     ylab="CovRatio",
     xlab="Observation Number",
     main="CovRatio")
abline(h = cutoff_pos, col ="red")
abline(h = cutoff_neg, col ="red")
text(x=1:length(covr)+1, y = covr, labels = ifelse(covr>cutoff_pos,names(covr),""),col="red")
text(x=1:length(covr)+1, y = covr, labels = ifelse(covr<cutoff_neg,names(covr),""),col="red")
cutlis <- list()

#Boxcox
boxcox <- boxcox(fit, lambda = seq(-1, 2, 1/10), plotit = TRUE,
                 eps = 1/50, xlab = expression(lambda),
                 ylab = "log–Likelihood")