**LUT School of Business and Management**
Bachelor's thesis, Business Administration
Strategic Finance

# The role of sustainability, manager tenure, and fund size in European mutual growth fund performance

# Vastuullisuuden, salkunhoitajan toimikauden pituuden ja rahaston koon rooli eurooppalaisten kasvusijoitusrahastojen suoriutumisessa

31.5.2021
Author: Fanni Welling
Supervisor: Jan Stoklasa

## ABSTRACT


**Author:** Fanni Welling


**Title:** The role of sustainability, manager tenure, and fund size in European mutual growth fund performance.


**School:** School of Business and Management

**Degree program:** Business Administration, Strategic Finance

**Supervisor:** Jan Stoklasa



**Keywords:** Mutual funds, fund characteristics, performance, fsQCA, sustainability, ESG



The aim of this thesis is to investigate the role of sustainability, manager tenure, and fund size on European mutual growth fund performance. While the relationships between fund characteristics and fund performance have been studied extensively, the literature review shows inconclusive evidence in the area. Methodologically, a novel approach in fund performance evaluation is made by using fuzzy set comparative qualitative analysis (fsQCA) and its enhancements. To analyze the above-mentioned possible relationships with fsQCA, linguistic rules are first formed based on findings from the literature review. These rules are then employed to recognize possible relationships in a sample of 429 European mutual growth funds in a period from March 2018 to March 2021.

The results do not indicate a strong relationship between manager tenure or fund size and risk-adjusted returns. Similarly, there is no significant evidence of a relationship between high sustainability, measured with Morningstar Sustainability ratings, and high risk-adjusted returns. However, although investing in funds with high Morningstar Sustainability ratings may not directly create abnormal financial performance, the findings suggest that high sustainability could help a fund avoid poor financial performance.

**TIIVISTELMÄ**

| | |
|---|---|
| **Tekijä:** | Fanni Welling |
| **Tutkielman nimi:** | Vastuullisuuden, salkunhoitajan toimikauden pituuden ja rahaston koon rooli eurooppalaisten kasvusijoitusrahastojen suoriutumisessa |
| **Akateeminen yksikkö:** | LUT-kauppakorkeakoulu |
| **Koulutusohjelma:** | Kauppatieteet, Strateginen rahoitus |
| **Ohjaaja:** | Jan Stoklasa |
| **Hakusanat:** | Sijoitusrahastot, rahastojen ominaisuudet, suoriutuminen, fsQCA, vastuullisuus, ESG |

Tämän kandidaatintutkielman tavoitteena on tutkia vastuullisuuden, salkunhoitajan toimikauden pituuden ja rahaston koon vaikutusta eurooppalaisten kasvusijoitusrahastojen suoriutumiseen. Vaikka rahastojen ominaisuuksien ja rahaston suoriutumisen välistä yhteyttä on tutkittu laajasti, aihealueen aiemmat tulokset ovat ristiriitaisia. Tutkielmassa hyödynnetään sumean logiikan kvalitatiivista vertailevaa analyysia (*engl. fuzzy set qualitative comparative analysis*, fsQCA) sekä sen uusia täydentäviä menetelmiä. Jotta vaikutuksia voidaan tutkia fsQCA:n avulla, muodostetaan ensin kirjallisuuskatsauksen perusteella lingvistisiä sääntöjä, joita hyödynnetään mahdollisten yhteyksien tunnistamiseen otoksessa. Tutkimuksen kohteena on 429 eurooppalaista kasvusijoitusrahastoa ajanjaksolla maaliskuusta 2018 maaliskuuhun 2021.

Rahaston koolla tai rahastonhoitajan toimikauden pituudella ei havaita olevan merkitsevää vaikutusta rahastojen riskikorjattuihin tuottoihin. Myöskään rahaston korkean Morningstar-vastuullisuusluokituksen ja korkeiden riskikorjattujen tuottojen välillä ei tulosten mukaan ole voimakasta yhteyttä. Vaikka rahaston korkean vastuullisuuden ei löydetty suoranaisesti johtavan rahastojen korkeisiin epänormaaleihin riskikorjattuihin tuottoihin, tulokset osoittavat rahaston korkean vastuullisuuden voivan auttaa rahastoa välttämään heikkoa taloudellista suoriutumista.

**TABLE OF CONTENTS**


**1 INTRODUCTION .................................................................................... 1**
1.1 Outline of the thesis .................................................................................... 3
1.2 Methodology .................................................................................... 3
1.3 Research objectives .................................................................................... 4
1.4 Thesis structure .................................................................................... 5

**2 LITERATURE REVIEW ON MUTUAL FUND CHARACTERISTICS .................. 6**
2.1 The effect of manager tenure on fund performance .......................................... 7
2.2 The effect of fund size on mutual fund performance ....................................... 10
2.4 Sustainable investing .................................................................................... 13
2.4.1 Measures of sustainable investing .................................................................. 13
2.4.2 Sustainability in mutual fund performance .................................................... 14
2.5 Conclusions from previous literature and hypotheses ................................ 18

**3 MEASURING FINANCIAL PERFORMANCE ..................................................... 20**
3.1 The Capital Asset Pricing Model (CAPM) ....................................................... 20
3.2 Fund performance measurements .................................................................. 21
3.2.1 Jensen's Alpha .................................................................................... 21
3.2.2 The Sharpe Ratio .................................................................................... 22

**4 ANALYSIS .................................................................................... 23**
4.1 Data description .................................................................................... 23
4.2 Methodology .................................................................................... 27
4.3 Process .................................................................................... 29
4.3.1 Linguistic variables .................................................................................... 30
4.3.2 Hypothesis demonstration and linguistic rules ............................................ 34

**5 RESULTS .................................................................................... 35**
5.1 Large fund size and fund performance ............................................................ 35
5.2 Long manager tenure and fund performance .................................................. 38
5.3 High sustainability and fund performance ...................................................... 40

**6 CONCLUSIONS .................................................................................... 43**

**References .................................................................................... 46**


**LIST OF FIGURES**


**FIGURE 1** Demonstration of fuzzy numbers for linguistic variables Low Alpha, Middle Alpha and High Alpha

**FIGURE 2** Demonstration of fuzzy numbers for linguistic variables Low Sharpe, Middle Sharpe and High Sharpe

**FIGURE 3** Demonstration of fuzzy numbers for linguistic variables Small size, Middle size and Large size

**FIGURE 4** Demonstration of fuzzy numbers for linguistic variables Short tenure, Middle tenure and Long tenure

**FIGURE 5** Demonstration of fuzzy numbers for linguistic variables Low sustainability and High sustainability

## LIST OF TABLES

**TABLE 1** Findings from past literature on the relationship between manager tenure and fund performance.

**TABLE 2** Findings from past literature on the relationship between fund size and fund performance.

**TABLE 3** Findings from past literature on the relationship between sustainability and fund performance.

**TABLE 4** Previous findings on the relationship between fund performance and selected fund characteristics.

**TABLE 5** Summary statistics for fund size and manager tenure by quartiles.

**TABLE 6** Summary statistics for sustainability by Morningstar sustainability ratings.

**TABLE 7** The hypothesis demonstrations with linguistic variables for fund size, manager tenure, and sustainability with risk-adjusted returns.

**TABLE 8** Results of the evaluation of the validity of rules large_size ⇒ low_sharpe and large_size ⇒ low_alpha

**TABLE 9** Results of the evaluation of the validity of rules large_size ⇒ high_sharpe and large_size ⇒ high_alpha

**TABLE 10** Results of the evaluation of the validity of rules long_tenure ⇒ high_sharpe and long_tenure ⇒ high_alpha

**TABLE 11** Results of the evaluation of the validity of rules Long_tenure ⇒ low_sharpe and long_tenure ⇒ low_alpha

**TABLE 12** Results of the evaluation of the validity of rules high_sustainability ⇒ high_sharpe and high_sustainability ⇒ high_alpha

**TABLE 13** Results of the evaluation of the validity of rules high_sustainability ⇒ low_sharpe and high_sustainability ⇒ low_alpha

## LIST OF MATHEMATICAL SYMBOLS AND ABBREVIATIONS

| | |
|---|---|
| FsQCA | Fuzzy set qualitative comparative Analysis |
| AUM | Assets under management |
| A, B | Fuzzy sets |
| $A \Rightarrow B$ | A set relation from A to B |
| $U$ | Universe of discourse |
| $\mu_A$ | Membership function of fuzzy set A |
| $Card(A)$ | Cardinality of fuzzy set A |
| $Supp(A \Rightarrow B)$ | Support of fuzzy set relation from A to B |
| $Disp(A \Rightarrow B)$ | Disproof of fuzzy set relation from A to B |

# 1 INTRODUCTION

Mutual funds are investment vehicles that pool investors' money to invest a collective amount in financial markets accordingly to the fund's specified objectives (Statista 2021). Mutual funds have been an increasingly popular investment vehicle since the establishment of the first one in the 1920s (Brooks & Tompkins 2002), all the way to over 122,500 mutual funds in 2019 (Statista 2020). The benefit of mutual funds is that they provide an opportunity for investment with professional know-how and provide an investor with risk diversification with the task of fund allocation outsourced to the fund manager (Cuthbertson, Nitzsche & O'Sullivan 2016; Kaur 2018).

This bachelor's thesis examines the relationship between selected mutual fund features and mutual fund performance. Recognizing the relationships between mutual fund characteristics and fund performance can help investors make conscious investment decisions (Yin-Ching & Mao-Wei 2003), and the relation between fund characteristics and fund performance has received attention in past studies (Cuthbertson et al., 2016). In this thesis, a fuzzy-set qualitative comparative analysis (fsQCA) approach is used to validate findings from previous studies, and to build on these earlier results. As a method, FsQCA comes with multiple upsides, such as overcoming the supposition of linear relationships between fund returns and other examined variables (Graham, Lassala & Ribeiro Navarrete 2020).

The raised concerns over the effect of climate change and the outcome of pandemics have attracted more attention to environmental and social risks. In this respect, sustainable investing forms have increased extensively during the past decade due to a rising demand of investors to reflect sustainability concerns in their investment decisions. (OECD 2020) Sustainable investing refers to an investment approach that considers environmental, social, and governance (ESG) criteria in portfolio selection and management to create both long-term competitive financial performance and positive societal impact. Due to various investment approaches within the field of sustainable investing, it is often labeled for example as green investing, responsible investing, or socially responsible investing (SRI). (US SIF n.d.) The use of ESG approaches has been driven by investor's increased demand to take more nonfinancial information into account in asset allocation decisions (OECD 2020). According to the Global Sustainable Investing Alliance review (2018), from 2016 to 2018, the amount of sustainable investing assets grew by 34 percent globally, which underlines the upward popularity of sustainable investing.

Although interest in ESG investing has grown over the past years among institutional investors, some are concerned that including ESG factors in their investments would cost them as lower risk-adjusted returns (Nagy, Kassam & Lee 2016). Past literature exposes a broad range of results and approaches on the impact of ESG on investment performance. There is a significant amount of variation in the findings of the existing literature; some show a positive impact of ESG criteria on overall investment performance, while others do not find any considerable effects. (OECD 2020) The inconsistency in past results and the absence of consensus in the subject might be due to the lack of standardized ways to measure sustainability (Dolvin, Fulkerson & Krukover 2019).

The three largest centers of sustainable investing are Europe, the United States, and Japan (OECD 2020). In Europe, the portion of total assets devoted to sustainable and responsible investment strategies of the overall market was 49 percent in 2018. The share was 26 percent in the US and 18 percent in Japan. Europe remains the most significant region for sustainable investors, with over 12 trillion euros out of global over 25 trillion euros devoted to sustainable investment strategies. (Global Sustainable Investment Alliance 2018) The high level of sustainable investments in Europe may propose sustainable investments yielding competitive returns in the region (Ibikunle & Steffen 2017). Therefore, in this thesis, a closer investigation into the performance of European mutual funds and the relationship between a fund’s degree of sustainability and performance is conducted. In addition, other fund characteristics’ effect on European mutual fund performance is examined. These characteristics are further identified after identifying research problems from a review of past literature.

### *1.1 Outline of the thesis*

Because this thesis will mainly focus on investigating the role of sustainability on mutual fund performance, the study is geographically outlined to Europe to see whether European sustainable funds could create competitive financial returns, as Ibikunle and Steffen (2017) stated as a possibility. The common region of the examined mutual funds also ensures that the funds are as comparable as possible in terms of regional regulation and the region's common currency.  Additionally, this will ensure that the market interest rate in the region is uniform. The research will consider only growth funds, i.e., funds that do not pay dividends, for optimal comparability of the funds' performance.

The data utilized in this study are retrieved from Morningstar Mutual Fund Screener. To diminish the number of changes of a fund and survivorship bias, cross-sectional data from a shorter period can be utilized (Brown, Goetzmann, Ibbotson & Ross 1992). By using three-year values from March 2018 to March 2021, the changes in fund management and liquidation are minimized. Furthermore, the Morningstar Sustainability Rating was announced in late 2016, which rules out the possibility of investigating the long-term role of sustainability. Due to the low amount of other than equity funds in the database, this study considers only mutual equity funds, i. e funds that allocate their investments mainly in the stock market. In the Morningstar Mutual Fund Screener, in April 2021, approximately 82 percent of European mutual funds were equity funds, which confirms the high number of mutual equity funds in the area.

### *1.2 Methodology*

To examine the effect of selected fund characteristics on fund performance, this study involves the following procedures: a literature review, recognizing research problems and forming hypotheses, data gathering, a fuzzy set qualitative comparative analysis (fsQCA), discussion of the empirical results, and drawing the conclusions.

The study begins with a literature review on selected topics, after which the research problems and hypotheses are formulated. Next, data are collected. As discussed earlier, in this thesis, secondary mutual fund data from Morningstar Mutual Fund Screener are used. Then, the sample is analyzed, and the identified hypotheses are tested by using a fuzzy set qualitative comparative analysis (fsQCA).

FsQCA was initially developed by Ragin (2008) based on the fuzzy set theory by Zadeh (1965). To our knowledge, fsQCA has been used only by Graham, Lassala, and Ribeiro Navarrete (2019, 2020) to study the relationship between mutual fund characteristics and mutual fund performance. Additionally, this thesis applies enhancements of fsQCA developed by Stoklasa, Luukka and Talášek (2017) and Stoklasa, Talášek and Luukka (2018), which have not been employed before to study mutual fund characteristics' role in mutual fund performance.

## *1.3 Research objectives*

The objective of this thesis is to examine if sustainability or other selected characteristics have an effect on the performance of European mutual funds, and thus provide information on whether investors seeking financial performance should aim to focus investments in funds with specific characteristics. The aim is to validate findings of past literature or, alternatively, discover new findings on the possible relationships using a novel method that has not been employed in performance evaluation before.

To fulfill the objectives of examining new approaches of fsQCA, other mutual fund characteristics in addition to the rate of sustainability are examined. As mentioned before, additional tested characteristics are identified from the past literature, and the sub-research questions are identified after the literature review in chapter 2.5. The main research question is as follows:

> *Is there a relationship between selected mutual fund characteristics and the performance measured with risk-adjusted returns of mutual growth funds registered in Europe?*

### *1.4 Thesis structure*

In terms of structure, the rest of the thesis follows the subsequent form: in the second chapter, a review of the past literature of fund characteristics is provided, along with the sub-research questions and hypothesis of the study. In the third chapter, fundamental theory, along with performance measurements that are applied to evaluate fund performance are explained. In chapter four, analysis is performed with data description, the definition of the methodology (fsQCA), and the analysis process. Chapter five concludes the results of the empirical analysis together with a discussion of them. Lastly, in chapter six, conclusions and limitations along with future research suggestions are provided.

# 2 LITERATURE REVIEW ON MUTUAL FUND CHARACTERISTICS

Mutual fund performance is both a popular and an important finance topic since funds' positive risk-adjusted returns have an association with market efficiency (Golec 1996). A significant benefit in mutual fund investing is risk diversification. The Modern Portfolio Theory (MPT) is a broadly applied financial theory presented by Harry Markowitz (1952). The theory demonstrates that a reasonable investor can maximize the expected return of a portfolio while minimizing market risk through diversification. According to the theory, an investment portfolio's risk carries two types of risk: systematic and unsystematic risk (Markowitz 1952; Sharpe 1964). Systematic risk is fixed in the whole market's volatility, while unsystematic risk represents an individual security's risk. Through diversification, an investor can construct portfolios so that another individual security's unsystematic risk offsets the unsystematic risk of individual security. (Barnett & Salomon 2006) Funds provide an investor the benefits of diversification with a minimum asset allocation (Kaur 2018; Cuthbertson et al., 2016).

Mutual funds can be categorized by their management style to actively managed mutual funds and passively managed mutual funds, also referred to as index funds. In the first case, funds aim to outperform the benchmark index, whereas in the latter case, funds seek to mimic a predefined benchmark index (Ferrari 2016, 90). For passively managed index funds, the returns are fairly similar to those of the index, and fund fees are typically lower than average since the fund mirrors the index, and hence the management costs are lower. For actively managed funds, the portfolio manager aims to outperform an investment benchmark index, such as the S&P 500 Index, and thus makes specific decisions on the holdings. To outperform the market, actively managed fund managers seek to defeat winner stocks producing abnormal performance or try to succeed in timing the market. Market timing stands for the fund manager's ability to forecast and employ foreseen movements in the market. (Friis & Smit 2004) Stock picking and market timing skills are more necessary in actively managed funds than in passively managed index funds that mimic market indices (Ejara and Nag 2011). Actively managed funds require more research compared to index funds and are thus more expensive, which typically reflects as more expensive shareholder fees (Tufano & Sevick 1997).

Various studies from the 60s (see e.g., Jensen 1968) have discussed whether fund managers can add financial value to a fund or if the possible outperformance of a fund is just pure luck. While some studies have found slight persistence of performance in actively managed mutual funds, most studies conclude that they cannot outperform the market (Ferreira, Keswani, Miguel & Ramos 2012). The persistence of performance determines how likely an outperforming fund or fund manager will continue to outperform the market in the future (Fortin, Michelson & James Jordan-Wagner 1999; Friis & Smit 2004). Hendricks, Patel, and Zeckhauser (1993) found evidence of short-term persistence in mutual fund performance, especially within poorly performing funds. Carhart (1997) found slight proof consistent with the hypothesis of skilled mutual fund managers but stated that almost all persistence of performance might be due to the one-year momentum effect with a lot of the enduring persistence of performance attributable to the worst-performing funds.

## *2.1 The effect of manager tenure on fund performance*

Although mutual funds have stated objectives of investment, fund managers often have an impact on the selection of individual securities when making investment decisions based on their abilities and risk preferences (Golec 1996; Fortin, Michelson & James Jordan-Wagner 1999). For this reason, investors look for skillful fund managers (Friis & Smit 2004). Manager tenure refers to the number of years the fund manager has been managing the fund and can be used as a proxy for managerial experience. As a measure of managerial experience, manager tenure could influence mutual fund performance. (Graham et al., 2019) Moreover, manager tenure evaluates the fund manager's survivorship at the position. A fund manager's long tenure can refer to the investment management company's satisfaction with the fund manager's skills and performance. However, it could also imply that the fund manager does not have better opportunities due to specialized skills or unimpressive performance in the past. (Golec 1996) In this thesis, manager tenure is measured by the number of years the current fund manager has been in his/her position, as in Morningstar (2021).

Some studies show that manager tenure does not affect fund performance (e.g., Fortin, Michelson & Jordan-Wagner 1999; Costa, Jakob & Porter 2006). Costa et al. (2006) examined the effect of market trends and manager tenure on risk-adjusted returns measured with the Carhart four-factor model. Their sample consisted of 1249 U.S. mutual equity funds' 36-month returns over the period 1990-2001. According to the results, they suggest no clear relationship between manager tenure and risk-adjusted performance. (Costa et al., 2006) Fortin et al. (1999) studied 800 equity and bond mutual funds over a ten-year period of 1985-1995 and did not find a relationship between manager tenure and fund performance. They did find a significant adverse relation between manager tenure and fund turnover, as well as a moderate positive relation between manager tenure and fund size. (Fortin et al., 1999) Brooks and Tompkins (2002) examined 474 mutual funds and found a slight negative link between manager tenure and mutual funds' risk-adjusted return measured with M-squared.

Golec (1996) argues that fund performance and fund manager tenure are significantly positively related, and better risk-adjusted performance could be expected from relatively young fund managers with a reasonably long (more than seven years) manager tenure. The study suggests that investors trying to achieve high yield should avoid funds with high management fees and favor large-sized funds managed by long-tenured managers. (Golec 1996) Lemak and Satish (1996) observed longer-tenured fund managers outperforming shorter-tenured fund managers, and that longer-tenured fund managers’ portfolios hold lower risk. Likewise, Filbeck and Tompkins (2004) found evidence of longer-tenured managers providing better risk-adjusted returns than other managers. They also show that longer-tenured managers charge lower fees than other managers and thus can operate more efficiently. (Filbeck & Tompkins 2004) Kjetsaa and Kieff (2016) examined a sample of blend funds from Morningstar and found a slight positive relation between manager tenure and fund returns. However, they point out that the results were not that robust because the data are imperfect since a manager's management performance records from previous funds are unavailable in the data from Morningstar. (Kjetsaa & Kieff 2016)

**Table 1 Findings from past literature on the relationship between manager tenure and fund performance**

| Author | Objectives | Data sample | Methodology | Results |
|---|---|---|---|---|
| **Brooks and Tompkins (2002)** | Investigating the effect of mutual fund characteristics on mutual fund performance. | Period 1989-1999; a sample of 474 mutual funds; geographically not specified. | A two-tailed Z-test and regression analysis. M-squared as a measure of risk-adjusted performance. | There is a slight adverse relationship between manager tenure and risk-adjusted returns. |
| **Costa, Jakob & Porter (2006)** | Examining how market trends and fund managerial experience affect the ability to outperform the market. | Period 1990-2001; 1249 mutual equity funds from the U.S. | Regression analysis. Alpha from a four-factor model as a performance measurement. | Longer-tenured managers do not outperform shorter tenured managers. |
| **Filbeck and Tompkins (2004)** | Investigating if there is a relation between manager tenure and risk-adjusted returns. | Period 1990-2000; sample size or geographical area not specified. | Regression analysis. M-squared as a measure of risk-adjusted performance. | Longer-tenured managers outperformed the market more than shorter tenured managers. Long-tenured managers were able to manage funds on lower expenses and thus more efficiently. |
| **Fortin et al. (1999)** | Researching how manager tenure affects mutual fund performance across all investment classes. | Period 1985-1995; 800 bond and equity funds; geographically not specified. | Comparison of short-term and long-term fund managers' performance and regression analysis. Alpha as a performance measurement. | Manager tenure does not affect mutual fund performance. There is an adverse relation between manager tenure and fund turnover. |
| **Golec (1996)** | To study if mutual fund manager's features affect fund fees, performance and risks. | Period 1988-1990; 530 mutual funds; geographically not specified. | A three-stage least squares (3SLS) regression analysis. Yield and Jensen's Alpha as performance measurements. | There is a positive connection between manager tenure and fund performance. |

| | | | | |
|---|---|---|---|---|
| **Kjetsaa and Kieff (2016)** | Exploring the effect of manager tenure, expenses and turnover on blend fund performance. | Period of 2002-2012; 559 blend funds; geographically not specified. | Regression analysis for three time horizons (3, 5 and 10 years). Returns as performance measurements. | There is a positive relation between manager tenure and mutual fund returns. |
| **Lemak and Satish (1996)** | Examining the differences in mutual fund performance and risk between longer-tenured mutual fund managers (> 10 years) and shorter tenured managers (< 10 years). | Period 1984-1994; 313 mutual funds; geographically not specified. | Comparison of short-term and long-term fund managers' performance. A regression analysis. Return as a performance measurement. | Longer-tenured (10 years or more) fund managers performed better than shorter tenured managers. |

Past literature for the relationship between fund manager tenure and mutual fund performance is concluded in Table 1. Much of the past literature (e.g., Golec 1996; Lemak & Satish 1996; Filbeck & Tompkins 2004; Kjetsaa & Kieff 2016) suggests there is a positive relationship between fund manager tenure and fund performance. These results suggest that experience could be a factor to consider when making investment decisions.

## *2.2 The effect of fund size on mutual fund performance*

Fund size, usually measured by assets under management (AUM), often affects the fund's management (Graham et al., 2019). Assets under management refers to the number of assets a fund manages on behalf of investors. According to Golec (1996), the amount of assets under management determines a fund's acceptance in the market, economies of scale, and past growth. On the other hand, some argue that as a fund's asset base grows, a fund manager's task to create value-added becomes harder, and for this reason, successful fund managers close funds from new money (Beckers & Vaughan 2001).

Multiple research (e.g., Golec 1996, Tufano & Sevick 1997) argue that bigger funds achieve cost advantages and thus show a positive relationship between fund size and performance. According to some, bigger funds can obtain cost advantages from fund expenses such as brokerage commissions, research costs, and administrative and overhead expenses. In support of this argument, according to Golec (1996), larger funds (over 280 million dollars under management) have economies of scale on expenses and fees and keep the fund's beta coefficient up and residual return variance down, thus improving fund survival and growth. Similarly, Tufano and Sevick (1997) found evidence of economies of scale with bigger fund size being inversely related to fund fees.

However, it has been presented that the large size of a fund could lead to reducing return expectations because funds with more assets are unable to invest as freely as smaller funds (e.g., Perold & Salomon 1991; Beckers & Vaughan 2001; Chen, Hong, Huang & Kubik 2004; Yan 2008; Chan, Faff, Gallagher & Looi 2009). Supporting this proposal, Perold and Salomon (1991) found diseconomies of scale for large funds because of the increased costs associated with larger transactions and as large trades are more difficult to implement. Similarly, Beckers and Vaughan (2001) propose that larger funds lose their flexibility since trading takes longer and opportunities vanish with the delay. Chen et al. (2004) found evidence that a large fund size has an adverse impact on fund performance. They do not find reasons such as heterogeneity in fund styles, possible correlations between other fund characteristics, or survivorship bias to explain this relationship. Instead, they find that fund size reduces fund performance mainly within funds that feature small-cap stocks. (Chen et al., 2004) Yan (2008) discovered similar results supporting the inverse relationship between fund size and performance. The research highlighted the significant role of liquidity in explaining this negative relation as it is stronger within funds with less liquid portfolios. (Yan 2008) Chan et al. (2009) examined a sample of Australian equity funds and argue that fund size is negatively associated with fund size.

Conclusions of past literature on the impact of fund size on fund performance are presented in Table 2. Based on past literature, it seems like small, nimble players might perform better than funds with a large asset base. Research (e.g., Yan 2008) emphasizes liquidity as a key feature to the success of smaller funds since they are able to liquidate investments and react to emerging opportunities. Although the relationship between fund size and fund performance has been widely studied, past research is inconclusive (Graham et al., 2019).

**Table 2 Findings from past literature on the relationship between fund size and fund performance**

| Author | Objectives | Data sample | Methodology | Results |
|---|---|---|---|---|
| **Beckers and Vaughan (2001)** | Examining how fund size affects investment performance | Period 1996-1999; 250 stocks from an Australian Index; Daily prices and trading volumes | Historical real-life simulation | Bigger funds are less flexible in implementing their ideas and thus creating value-added is harder as the number of assets under management grow. |
| **Chan et al. (2009)** | Investigating if fund size affects performance. Identifying the causes for the possible relation. | Period 1998-2001 (40mths); 35 Australian equity funds | Regression analysis and a simulation method. | Fund size lowers performance, especially for funds with highly active trading approaches. |
| **Chen et al. (2004)** | To investigate if fund size affects fund performance. | Period 1962-1999; 3439 funds from the U.S. | Regression analysis. Performance measured with CAPM model, three- and four-factor models. | A negative relationship between fund size and fund performance mainly caused by the lack of liquidity. |
| **Golec (1996)** | To study if mutual fund manager's features affect fund fees, performance and risks. Also the effect of fund size is examined. | Period 1988-1990; 530 mutual funds; geographically not specified. | Regression analysis. Alpha and yield as performance measures. | Larger funds discover economies of scale. Large funds' fees are lower leading to larger yields. |
| **Perold and Salomon (1991)** | To detect the right amount of assets under management for financial maximization. | Examples from T.Loeb's (1983) article "Trading cost: The Critical Link Between Investment Information and Results." Period 1982; 1200 observations. | A mathematical analysis using a wealth-maximizing tradeoff. Alpha as a performance measure. | The optimal fund size is when trading costs exceed the opportunity cost of not trading. A larger asset base than that leads to higher opportunity costs and lower returns. |

| | | | | |
|---|---|---|---|---|
| **Tufano and Sevick (1997)** | Researching the relationship between fund board structure and fund fees. Also the relationship between fund size and fees is examined. | Period 1991-1992 (12mths); 1587 U.S. open-end mutual funds. | Regression analysis. | Fund fees are inversely related to fund size, and thus larger funds have economies of scale. |
| **Yan (2008)** | To examine the impact of liquidity and investment style on the relationship between fund size and fund performance. | Period 1993-2002; 1024 actively managed U.S. mutual funds. | Cross-sectional regression analysis and a portfolio approach. Performance measured with Alpha, CAPM model, three- and four-factor models. | A negative relationship between fund size and fund performance. Liquidity is proposed as an important reason to cause this relation. |

## *2.4 Sustainable investing*

The Global Sustainable Investing Alliance describes sustainable investing as "an investing approach that considers environmental, social and governance (ESG) factors in portfolio selection and management." In their review (GSIA 2018), they use a general definition of sustainable investing and do not draw differences between sustainability and associated terms, such as socially responsible investing (SRI) and responsible investing. Similarly, for the sake of clarity, this thesis mainly uses the term sustainable investing, referring to SRI investing or other investments complying with ESG criteria.

### *2.4.1 Measures of sustainable investing*

There seem to be controversial suggestions for the definition of sustainable investing, and the lack of consensus in socially responsible investing might be partially due to the absence of a broadly accepted way to define it. The absence of standardization and common opacity could explain the moderately inconsistent results of existing research on sustainable investment performance. (Dolvin, Fulkerson & Krukover 2019) The industry's first general sustainability rating for funds was published by Morningstar in 2016. Morningstar is an investment research company that gathers and analyzes fund, stock, and general market data (Morningstar 2020), and is the accepted leader as a mutual fund data provider (Dolvin et al., 2019). The Morningstar Sustainability Rating helps investors evaluate approximately 20,000 mutual and exchange-traded funds (ETF) by their sustainability. The rating aims to offer a reliable and objective way for investors to see how portfolios meet environmental, social, and governmental (ESG) challenges. (Morningstar 2016)

The Sustainability Rating uses ESG data from Sustainalytics, an ESG and corporate governance research company owned by Morningstar (Morningstar 2020). The rating is a measure of a fund's adherence to ESG factors and classifies each fund to a category between one globe (low sustainability) and five globes (high sustainability). Funds in the top 10 percent receive a Sustainability Rating of five globes, while the bottom 10 percent are categorized with one globe. The Morningstar Sustainability rating is based on a portfolio's Morningstar Historical Portfolio Scores. Morningstar Historical Portfolio Sustainability Score is a weighted average of 12 months of Morningstar Portfolio Sustainability Scores, which is an asset-weighted average on Sustainalytics' company-level ESG Risk Rating. The company-level ESG Risk Rating is a measure of the extent up to which a company's economic value can be risk-driven by ESG challenges. To have a Morningstar Sustainability Rating, at least 50 percent of a fund's assets must be covered by company-level ESG scores from Sustainalytics. The score is updated each month accordingly to the data from Sustainalytics. (Morningstar 2019)

### *2.4.2 Sustainability in mutual fund performance*

It has been shown that socially responsible investors could be ready to decrease financial performance to invest by their social preferences (Riedl and Smeets 2017). However, there is a large amount of discussion on sustainable funds' performance and whether a sustainable investment can perform financially competitively. In this section, previous literature on the relationship between sustainability and mutual fund performance is presented.

Dolvin, Fulkerson, and Krukover (2019) analyzed the effect of Morningstar Sustainability scores on risk-adjusted returns of funds. They found no practical difference in risk-adjusted returns between funds with high sustainability ratings and conventional funds. However, the results point that most funds with high sustainability ratings confine to large-cap funds and feature an apparently higher risk along with a degree of weaker diversification. Hamilton, Jo, and Statman (1993) and Bello (2005) found similar results. Hamilton et al. (1993) found that the performance of socially responsible funds is not significantly different from conventional funds. Bello (2005) did not find a notable difference between the performance of socially responsible and conventional funds. Furthermore, the effect of diversification did not differ between the fund groups. (Bello 2005) Steen, Moussawi, and Gjolberg (2020) analyzed the relationship between the Morningstar Sustainability Rating and 146 Norwegian mutual funds and did not find any difference in risk-adjusted returns. However, due to geographical bias, they analyzed European categorized funds separately and found higher positive risk-adjusted returns for funds with high ESG.

Nagy, Kassam, and Lee (2016) analyzed two strategies built using ESG data from MSCI. The back-tested models were the *"ESG tilt,"* which overweighs stocks with higher ESG ratings, and the *"ESG momentum,"* which overweighs stocks that have upgraded their ESG rating recently. As a result, they found that both model portfolios outperformed the global benchmark index. Furthermore, both portfolios improved their ESG profile during the eight years. (Nagy et al., 2016) Henke (2016) compares the financial effect of ESG criteria on socially responsible bond funds and their conventional pairs in the US and Euro area during 2001 – 2014 and found that socially responsible bond funds outperformed their conventional fund match annually. Nofsinger and Varma (2014) compared socially responsible mutual funds to their matched conventional funds and found socially responsible funds to outperform conventional funds during periods of market crisis and, surprisingly, to underperform through non-crisis periods. They propose that investors could value the asymmetry of conventional and socially responsible mutual funds for downside protection. (Nofsinger & Varma 2014)

**Table 3 Findings from past literature on the relationship between sustainability and fund performance**

| Author | Objectives | Data sample | Methodology | Results |
|---|---|---|---|---|
| **Bello (2005)** | To examine the effects of socially responsible investing on portfolio diversification and fund performance. | Period 1994-2001; 42 socially responsible funds provided by Morningstar and 84 conventional funds from the U.S. | Regression analysis. Comparing socially responsible with their conventional pairs. Performance measured with Jensen's Alpha, Sharpe Ratio and excess standard deviation adjusted return. | There is no notable difference between the performance or diversification of socially responsible and conventional funds. |
| **Dolvin et al. (2019)** | Investigating the effect of sustainable investing on investment performance. | Period 2012-2016; 1853 U.S. mutual funds. | Performance measured with Carhart alpha. Sustainability measured with the Morningstar Sustainability scores. | No difference in risk-adjusted returns between sustainable and conventional funds. However, sustainable funds limited to large-cap funds and thus can feature a higher risk and weaker diversification. |
| **Hamilton et al. (1993)** | Evaluating the financial effect of socially responsible investing in mutual fund performance. | Period Jan 1981 - Dec 1990; 32 socially responsible funds and 150 conventional funds. | Performance comparison between socially responsible and conventional funds. Jensen's Alpha as a performance measure. The selected funds were identified as socially responsible funds by their managers. | There is no practical difference between the performance of socially responsible and conventional funds. |
| **Henke (2016)** | To examine the financial effect of screening ESG criteria on corporate bond fund portfolios. | Period 2001-2014; 103 socially responsible and 309 matched conventional bond mutual funds from U.S. and Eurozone | Regression analysis. Comparing socially responsible funds with their conventional pairs. Performance measured with risk-adjusted returns (a five-factor model). Sustainability is measured with ESG ratings based on information provided by the US Sustainable Investment Forum and the European Social Investment Forum. | Socially responsible bond mutual funds performed better than their conventional pairs annually. |

| | | | | |
|---|---|---|---|---|
| **Nagy et al. (2016)** | To investigate if ESG factors of an investment affect investment performance. | Period 2007-2015; global MSCI stock data. | Back-testing two global model portfolios that regard ESG criteria: "ESG tilt" and "ESG momentum." Alpha as a performance measure. MSCI ESG ratings as a sustainability measure. | Both tested portfolios that consider ESG criteria beat the global benchmark index MSCI World Index. |
| **Nofsinger & Varma (2014)** | To examine the performance of socially responsible funds during periods of market crisis and periods of non-crisis. | Period 2000-2011; 240 U.S. equity mutual funds and their 209 conventional pairs | Regression analysis. CAPM, three-factor and four-factor models as performance measures. | Socially responsible mutual funds outperform their conventional pairs in periods of market crisis and underperform conventional funds during periods of non-crisis. |

Results from past literature are concluded in Table 3. Some previous literature presents that there is no relationship between environmental, social, and governmental ratings and fund performance. Alternatively, there might be a positive relationship between high sustainability and fund performance. Furthermore, taking the findings of Nofsinger et al. (2014) into account, there could be signs of funds with high sustainability outperforming conventional funds in the past year because of the financial effects of the global COVID19 pandemic. According to Boffo and Patalano (2020), at the beginning of the pandemic, sustainable market actors, including Bloomberg, Morningstar, and MSCI, observed ESG funds and indices outperforming standard investments by losing less value than traditional indices.

## 2.5 Conclusions from previous literature and hypotheses

The performance of mutual funds has been a popular and widely studied topic in finance. Reviewing the literature on the relationship between mutual fund features and mutual fund performance reveals some contradictory findings. In this chapter, past literature on the effect of manager tenure, fund size and sustainability on mutual fund performance was reviewed. Inconsistencies in past findings were found in all reviewed characteristics. Thus, it is reasonable to investigate all the three reviewed characteristics.

In chapter 1.3, the following main research question of this thesis was identified: *Is there a relationship between selected mutual fund characteristics and the performance measured with risk-adjusted returns of mutual growth funds registered in Europe?*

Based on the selected characteristics, sub-research questions are formed as follows:

1. *Is there a relationship between European mutual equity growth fund size measured with assets under management and their performance?*
2. *Is there a relationship between fund manager tenure and the performance of European mutual equity growth funds?*
3. *Is there a relationship between Morningstar Sustainability Rating and the performance of mutual equity growth funds registered in Europe?*

The main results from past literature are concluded in Table 4. It must be noted that the studies were methodologically different, and while some of the studies back-tested different sustainable investing strategies and compared them to their conventional pairs, some tested the effect of the characteristic on risk-adjusted returns by for example using linear regression analysis. Additionally, the sustainability approach, the performance measurements, market, and methodologies differed in the studies. The table and the signs concluded in it are only supposed to give an idea of the previous literature’s findings on the relationship between these characteristics and fund performance to form hypotheses for the sub-research questions.

**Table 4 Previous findings on the relationship between fund performance and selected fund characteristics**

In the table, past findings on the relationship between fund size, manager tenure, sustainability and fund size are presented. In the table, '+' indicates a positive relationship, '-' indicates a negative relationship, and '0' refers to no relationship.

| | Size | Manager tenure | Sustainability |
|---|---|---|---|
| Beckers and Vaughan (2001) | - | | |
| Bello (2005) | | | 0 |
| Brooks and Tompkins (2002) | | - | |
| Chan et al. (2009) | - | | |
| Chen et al. (2004) | - | | |
| Costa et al. (2006) | | 0 | |
| Dolvin et al. (2019) | | | 0 |
| Filbeck and Tompkins (2004) | | + | |
| Fortin et al. (1999) | | 0 | |
| Golec (1996) | + | + | |
| Hamilton et al. (1993) | | | 0 |
| Henke (2016) | | | + |
| Kjetsaa and Kieff (2016) | | + | |
| Lemak and Satish (1996) | | + | |
| Nagy et al. (2016) | | | + |
| Perold and Salomon (1991) | - | | |
| Steen et al. (2020) | | | + |
| Tufano and Sevick (1997) | + | | |
| Yan (2008) | - | | |
| Author's expected sign | - | + | + |

Based on past literature summarized in Table 4, we focus especially on the following hypotheses tested in the empirical section of this thesis:

$H_1$: *If fund size is large then risk-adjusted returns are low.*

$H_2$: *If manager tenure is high then risk-adjusted returns are high.*

$H_3$: *If Morningstar Sustainability Rating is high then risk-adjusted returns are high.*

# 3 MEASURING FINANCIAL PERFORMANCE

In this chapter, relevant theories and concepts that are applied in the analysis are introduced. Although this study will focus on fund characteristics' relationship with fund performance, it is useful to determine the fundamental finance theories such as the Capital Asset Pricing Model, from which many performance measures are stemmed from. The empirical section of this thesis will apply the performance measurements Jensen's Alpha and Sharpe Ratio to examine fund performance. Therefore, these measures are also presented in this section.

## *3.1 The Capital Asset Pricing Model (CAPM)*

The Capital Asset Pricing Model (CAPM) is a vital and much-employed model in investment performance literature from which different performance measures are often derived. The Capital Asset Pricing Model is developed on the Diversification and Portfolio theory by Markowitz (1952) and developed by Sharpe (1964) and Lintner (1965). CAPM estimates portfolio performance while adjusting the portfolio's risk level. CAPM considers only one risk factor, the market premium, i.e., the expected rate of return of the market portfolio. CAPM is an equilibrium model that deduces that portfolio return's covariance along with the return of the market portfolio to explain changes in excess portfolio returns. (Fama & French 2004) The Capital Asset Pricing Model formula is presented in equation 1.

$$E(r_i) = r_f + \beta_i\left[E(r_m) - r_f\right] \quad (1)$$

Where $E(r_i)$ is the expected rate of return of portfolio $i$, $r_f$ is the risk-free rate, $\beta_i$ is the beta coefficient of portfolio $i$, and $E(r_m)$ is the expected rate of return of the market portfolio.

The beta coefficient in CAPM presents the covariance risk of assets in the market relative to the average covariance risk of assets. (Fama & French 2004) Thus, the beta coefficient in CAPM is the volatility of the portfolio to the overall market. The market beta can be calculated as in equation 2.

$$\beta_i = \frac{\mathrm{cov}(r_i, r_m)}{\sigma^2(r_m)} \qquad (2)$$

Where $\beta_i$ is beta coefficient of portfolio $i$, $\mathrm{cov}(r_i, r_m)$ is the covariance of the return of the portfolio with the return of the market, and $\sigma^2(r_m)$ is the variance of the market return.

## *3.2 Fund performance measurements*

To analyze and compare fund performance over a certain period, one could always use fund returns. However, to understand fund returns thoroughly, the difference in risk levels should be considered to evaluate performance realistically. Because of the continually increasing popularity of the topic, there is a variety of models to measure financial performance (see e.g., Jensen 1968; Cahart 1997). This section reviews some of the most common models and risk-adjusted returns as performance measurements. These models are applied later in the study's empirical section as measures of the financial performance of the mutual funds.

### *3.2.1 Jensen's Alpha*

Alpha, also referred to as Jensen's Alpha or Jensen's Measure is a Capital Asset Pricing Model-based performance measurement derived by Michael Jensen (1968). Over the decades, alpha has received multiple meanings due to its many different variations and purposes (Barillas & Shanken 2017). In this thesis, the classic Jensen's Alpha is employed.

Alpha is a risk-adjusted measure of portfolio performance that indicates the abnormal return of a portfolio. Alpha is used to measure how much the portfolio's realized return varies from the expected return, as determined by CAPM. Alpha can be either positive, negative, or zero. In other words, if Alpha is positive, the portfolio outperforms the hypothetical return of the benchmark market portfolio with the same risk level. Alpha measures the portfolio's return attributive to the manager's skill or luck (Golec 1996). Jensen's Alpha can be calculated as follows:

$$\alpha_i = r_i - [\beta_i(r_m - r_f)] \tag{3}$$

Where $\alpha_i$ is the Alpha of portfolio $i$, $r_i$ is the return of portfolio $i$, $\beta_i$ is beta for portfolio $i$, $r_m$ is the market return and $r_f$ is the risk-free rate.

### *3.2.2 The Sharpe Ratio*

The Sharpe Ratio was developed by William Sharpe (1966). The ratio measures the amount of return received per unit of risk. Instead of using beta to measure risk as in Capital Asset Pricing Model-based performance measurements, the Sharpe Ratio uses a portfolio's standard deviation as a risk measurement. (Sossong 2014, 18) The Sharpe Ratio can be calculated as follows:

$$S_i = \frac{r_i - r_f}{\sigma_i} \tag{4}$$

Where $S_i$ is the Sharpe Ratio for portfolio $i$, $r_i$ is the return of portfolio $i$, $r_f$ is the risk-free rate and $\sigma_i$ is the standard deviation of portfolio $i$. Because the Sharpe Ratio compares the portfolio's return to the portfolio's risk, the higher the ratio, the more efficient the measured portfolio is (Wang, Chen, Lian & Chen 2020). However, if no other information about an investment is given, it cannot be estimated whether a Sharpe Ratio is good or not. Instead, the Sharpe Ratio should always be for comparison of an investment with other similar types of investments. (Morningstar 2015)

# 4 ANALYSIS

The study will be conducted using a fuzzy set qualitative comparative analysis by Ragin (2008) and its enhancements developed by Stoklasa et al. (2017, 2018). These methodologies are applied to examine if there is a relationship between mutual equity fund performance and the selected fund characteristics, namely, manager tenure, fund size measured with assets under management, and Morningstar sustainability rating. In this chapter, the description of data and methodology, as well as the description of the methodology process, is provided.

## *4.1 Data description*

The data used in this thesis was retrieved from Morningstar Mutual Fund Screener in March 2021. Morningstar provides global financial data for example for mutual funds, which is regularly used in studies. For example, Filbeck & Tompkins (2004), Golec (1996), Fortin, Michelson & James Jordan-Wagner (1999), and Kjetsaa & Kieff (2016) apply data from Morningstar in their research. Morningstar provides data for all examined characteristics: fund size measured with net assets under management, manager tenure, and the Morningstar Sustainability Rating. For manager tenure, the years the current fund manager has been in his/her position is provided in the data. If a fund is managed by a team, the tenure of the manager who has been in the position the longest is shown. If the fund has only one manager who has been in the position for less than six months, the tenure is not displayed in the data. (Morningstar Office 2021) Two different measures of risk-adjusted returns are used to evaluate fund performance: Jensen’s Alpha and Sharpe Ratio. Both measures are averages from the past three years’ average values from monthly returns.

The Morningstar Mutual Fund Screener holds data from over 31 000 mutual funds at the time of retrieval (March 2021). The following search criteria were used to select the funds for the sample:

1. *Europe Developed* and *Europe Developing* as the largest geographical regions. This criterion limited the funds to 3583, out of which 3378 were registered in Europe Developed and 205 in Europe Developing.

2. *Growth* as a fund distribution to exclude any dividend-paying funds from the study.
3. *Euro* as the currency. This is to ensure the best possible comparability of the funds.
4. A fund must be over three years old (March 2018 – March 2021) to have enough data for a three-year performance evaluation.
5. Funds must have a value of the Morningstar Sustainability Rating. This requires at least 50 percent of a fund’s assets to be covered by company-level ESG scores from Sustainalytics (Morningstar 2019).

Other than equity funds were eliminated from the sample due to the low number of other funds and their available variables. Overall, all suitable mutual funds were selected for the study accordingly with the search criteria presented above. The selected funds were then evaluated, and there appeared to be some possible duplicates that had, for example, an identical amount of assets under management and the same average market cap. Potential duplicates were eliminated by excluding funds with the same average market cap and assets under management to avoid bias in the data. Additionally, all possible funds with missing required values, such as assets under management, manager tenure, or Morningstar Sustainability Rating were excluded from the sample. After eliminating the duplicates and funds with missing values, the final sample consisted of 429 mutual growth equity funds registered in Europe.

Table 5 summarizes the main statistics for the examined data variables fund size and manager tenure by quartiles. For the whole sample, the mean 3-year annualized return is 8.11 percent, the mean 3-year Sharpe Ratio is 0.44, and the 3-year Jensen’s Alpha is -0.16. Overall, according to these statistics, the sample has underperformed the market on average since Jensen’s Alpha is negative (-0.16) and the Sharpe Ratio is relatively low (0.44). The largest funds with over 665 million euros under management have the highest annualized 3-year return (9.79 %), Alpha (0.51), and Sharpe Ratio (0.52). For manager tenure, 3-year return (9.12 %), Alpha (0.38), and Sharpe Ratio (0.48) are the highest for funds in quartile 3. These basic statistics suggest that managers who have been in their positions for 7.92-12.08 could provide more financial value than other managers.

**Table 5 Summary statistics for fund size and manager tenure by quartiles**

| | Quartiles | | | | All funds |
|---|---|---|---|---|---|
| | **1 (Bottom 25 %)** | **2** | **3** | **4 (Top 25 %)** | |
| **Fund size (AUM)** | | | | | |
| Number of funds | 107 | 108 | 107 | 107 | 429 |
| Max (M€) | 78.39 | 235.03 | 664.91 | 7124.65 | 7124.65 |
| Median (M€) | 37.86 | 135.12 | 415.99 | 1275.78 | 235.03 |
| Mean (M€) | 535.83 | 142.97 | 486.72 | 829.76 | 564.78 |
| 3-year return (annualize | 6.09 % | 8.16 % | 8.38 % | 9.79 % | 8.11 % |
| 3-year Sharpe Ratio | 0.36 | 0.44 | 0.46 | 0.52 | 0.44 |
| 3-year Jensen's Alpha | -0.47 | -0.50 | -0.18 | 0.51 | -0.16 |
| **Manager tenure (years)** | | | | | |
| Number of funds | 108 | 107 | 108 | 106 | 429 |
| Max (years) | 3.58 | 7.83 | 12.08 | 23.58 | 23.58 |
| Median (years) | 2.13 | 5.92 | 10.25 | 15.17 | 7.83 |
| Mean (years) | 2.05 | 5.73 | 10.05 | 15.98 | 8.42 |
| 3-year return (annualize | 8.38 % | 7.19 % | 9.12 % | 7.71 % | 8.11 % |
| 3-year Sharpe Ratio | 0.45 | 0.42 | 0.48 | 0.42 | 0.44 |
| 3-year Jensen's Alpha | -0.30 | -0.27 | 0.38 | -0.45 | -0.16 |

Table 6 presents the summary statistics for the sample sorted by Morningstar Sustainability Ratings. As mentioned before, Morningstar gives a rating of one to the lowest 10 percent of funds conforming to ESG factors, and, respectively, funds in the top 10 percent receive a rating of five (Morningstar 2019). In this sample, however, only 7 percent of funds have a rating of one, while 14 percent have a rating of five, meaning the sample holds more funds with high Sustainability Ratings than low Sustainability Ratings.

When observing the performance of the sample funds by their Morningstar Sustainability Ratings, it can be noticed that funds with the highest rating have the best values for the annualized 3-year return (10.90 %), 3-year Sharpe Ratio (0.59) and Jensen’s Alpha (0.81). The average values of these measures for funds with a rating of 1-3 are lower, and for example, the average Jensen’s Alpha is negative. The basic statistics presented in Table 6 would suggest funds with high Morningstar

Sustainability Ratings performing better than funds with low ratings and performance increasing with the ratings.

**Table 6 Summary statistics for sustainability by Morningstar Sustainability Ratings**

| | Morningstar Sustainability Rating | | | | | |
|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **All funds** |
| Number of funds | 30 | 75 | 174 | 91 | 59 | 429 |
| % of funds | 7 % | 17 % | 41 % | 21 % | 14 % | 100 % |
| 3-year return (annualized) | 3.87 % | 7.11 % | 8.11 % | 8.49 % | 10.90 % | 8.11 % |
| 3-year Sharpe Ratio | 0.28 | 0.37 | 0.44 | 0.48 | 0.59 | 0.44 |
| 3-year Jensen's Alpha | -1.28 | -0.90 | -0.36 | 0.58 | 0.81 | -0.16 |

The sample is not completely bias-free. Brown et al. (1992) pointed out survivorship bias in fund performance research, i.e., if funds that did not survive (dead funds) through the studied period are excluded from the sample, the performance can be upwardly biased. Because there is no data of dead funds, the database consists only of funds that survived through the studied period, and the sample is not free of survivorship bias. Also, some funds might at times choose not to release information about their performance. Second, as past scores for Morningstar Sustainability Rating are unavailable, and the data considers only present sustainability scores, the relationship between past performance and present score for Morningstar Sustainability rating is compared. In addition, as the Morningstar Sustainability Rating considers only funds with at least 50 percent of assets covered by company-level ESG scores, some funds with non-sustainable (or highly sustainable) assets are not included in the sample. Lastly, since manager tenure is measured with the years the current manager has been in his/her position and records from managing another company's fund are unavailable and omitted, manager tenure is not a robust proxy for the manager's experience. However, a longer tenure of a manager might imply the fund management company's satisfaction with the manager's past performance (Golec 1996).

### 4.2 Methodology

To evaluate the performance of the 429 mutual funds of the sample, fuzzy set qualitative comparative analysis (fsQCA) is used. Fuzzy set qualitative comparative analysis (fsQCA) was originally developed by Ragin (2008) based on fuzzy-set theory initially introduced by Zadeh (1965). In addition to the original fsQCA, the analytic approach in this thesis includes improvements of fsQCA by Stoklasa et al. (2017, 2018).

Before further definitions of fsQCA, the basic notions of fuzzy set theory by Zadeh (1965) are defined. Let a fuzzy set $A$ be defined on a non-empty set $U$ (universe of discourse, e.g., sample funds in this study) by a mapping $\mu_A : U \rightarrow [0,1]$, where $\mu_A$ is a membership function of a fuzzy set $A$. For each $x \in U$, the value $\mu_A(x) = A(x)$ is a degree of membership of $x$ in a fuzzy set $A$. A fuzzy set can include both quantitative and qualitative assessments; for example, in fuzzy sets, the qualitative term "fully in" is represented by 1 and "fully out" by 0. The values between 0 and 1 represent the partial membership of each observation $x$ to a fuzzy set. (Zadeh 1965) A set of only values 0 or 1 is called a crisp (real-valued) set (Zimmermann 2010).

The original fsQCA (F1) consistencies and coverages defined by Ragin (2008) are calculated as in equations (5) and (6).

$$Consistency_{F1}(A \Rightarrow B) = \frac{Card(A \cap B)}{Card(A)} = \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big)}{\sum_{i=1}^{n} A(x_i)} \tag{5}$$

$$Coverage_{F1}(A \Rightarrow B) = \frac{Card(A \cap B)}{Card(B)} = \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big)}{\sum_{i=1}^{n} B(x_i)} \tag{6}$$

Where $A$ and $B$ are fuzzy sets on a non-empty universal set $U = \{x_1, x_2, \dots, x_n\}$, $Card$ is the cardinality of a fuzzy set and $x_i$ is the $i_{th}$ observation in a fuzzy set. Here it is assumed that $Card(A) \neq 0$ and $Card(B) \neq 0$. For F1, the higher the consistency is, the more proof in favor of the evaluated rule there is. However, even if the consistency was high, if the coverage is low, the results are not convincing as there are not a lot of cases explaining the results. (Ragin 2008)

Later on, Stoklasa et al. (2017) proposed F2 consistency and coverage measures, which remove the effect of ambivalent evidence. F2 consistency and coverage measures can be calculated as in equations (7) and (8).

$$Consistency_{F2}(A \Rightarrow B) = \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big) - \min(A(x_i), B(x_i), \bar{B}(x_i))}{\sum_{i=1}^{n} A(x_i)} \quad (7)$$

$$Coverage_{F2}(A \Rightarrow B) = \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big) - \min(A(x_i), B(x_i), \bar{A}(x_i))}{\sum_{i=1}^{n} B(x_i)} \quad (8)$$

Where $\bar{B}(x) = 1 - B(x)$ for all $x \in U$, and $\bar{A}(x) = 1 - \mathrm{A(x)}$ for all $x \in U$. That is, $\bar{B}$ is $not\ B$ and $\bar{A}$ is $not\ A$. They also suggested F3 consistency and coverage measures which likewise reflect the ambivalent evidence and consider pure counterevidence of the examined relationship (Stoklasa et al., 2017). F3 measures can be calculated as follows:

$$Consistency_{F3}(A \Rightarrow B) = max\left\{0; \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big) - \min(A(x_i), \bar{B}(x_i))}{\sum_{i=1}^{n} A(x_i)}\right\} \quad (9)$$

$$Coverage_{F3}(A \Rightarrow B) = max\left\{0; \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big) - \min(B(x_i), \bar{A}(x_i))}{\sum_{i=1}^{n} B(x_i)}\right\} \quad (10)$$

Where $\bar{B}$ is $not\ B$ and $\bar{A}$ is $not\ A$. Shortly, Stoklasa et al. (2018) suggested an updated measure F4 for consistency and coverage measures. The F4 consistency and coverage measures are presented in equations (11) and (12).

$$Cons_{F4}(A \Rightarrow B) = \frac{1}{2}\ \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big) - \ \min(A(x_i), \overline{\mathrm{B}}(x_i)\ )}{\sum_{i=1}^{n} A(x_i)} \tag{11}$$

$$Cove_{F4}(A \Rightarrow B) = \ \frac{1}{2}\ \frac{\sum_{i=1}^{n} \min\big(A(x_i), B(x_i)\big) - \ \min(B(x_i), \overline{\mathrm{A}}(x_i)\ )}{\sum_{i=1}^{n} B(x_i)} \tag{12}$$

Where $\bar{B}$ is $not\ B$ and $\bar{A}$ is $not\ A$. In F4, if consistency is higher than 0.5, there is more evidence supporting the examined rule $A \Rightarrow B$ than there is supporting $A \Rightarrow not\ B$. That is, there is more evidence in favor of the relationship than against it in the sample. (Stoklasa et al., 2018) In F1, F2, and F3, high values are considered to have a similar interpretation.

In addition to the consistency and coverage measures, Stoklasa et al. (2017) proposed degree of support and degree of disproof that provide more perception concerning the support in favor or evidence against the examined relationship in the data. The measure enables detailed investigation of support and disproof of the given rule in the data sample. (Stoklasa et al., 2017)

### *4.3 Process*

The examination is conducted using Matlab R2020b. In the study, the aim is to examine how sustainability, fund size and manager tenure of a fund affect mutual fund performance. Thus, rules based on past literature reviewed in chapter 2 are tested to see if there is any pattern following these rules in the sample. Examining a rule $A \Rightarrow B$, fuzzy set $A$ is the input variable and, respectively, $B$ is the output variable. In this analysis, the input variables are manager tenure (years), fund size (net assets under management), and sustainability (Morningstar Sustainability Rating). The output variables are risk-adjusted return measures Sharpe Ratio and Jensen's alpha.

### *4.3.1 Linguistic variables*

First, to identify meaningful groups of cases, all conditions and outcomes should be calibrated using theoretical and substantive knowledge to the interval [0, 1], where 0 means fully out and 1 means fully in, and values between 0 and 1 indicate a partial membership in the group (Ragin 2008). Fuzzy sets can be created with linguistic labels such as low, middle, and high that describe a particular linguistic variable (Klir & Yuan 1995). Here, it is reasonable to form linguistic variables labeled low, middle, and high for Sharpe Ratio, Jensen's alpha, and Morningstar Sustainability Rating. For fund size, measured with net assets under management, variables are labeled small, middle, and large, and for manager tenure, the variables are labeled short, middle, and long. In addition to forming the variables, summary statistics of the input variables are presented.

First, the linguistic variables for output variables are formed. Jensen's Alphas over zero are preferred as they have then outperformed the market. Respectively, Alphas under zero are not preferred since they then have underperformed the market. There are no generally *good* values for Alpha other than positive Alphas because the preferred Alpha depends on an investor's preferences. Here, a possibility would be to use a crisp set with values negative [0] or positive [1] Alpha. However, it must be considered that an Alpha just over zero has outperformed the market slightly and thus may not be as preferable as a higher Alpha. Values of Alpha in the sample range from -12.44 to 14.23.

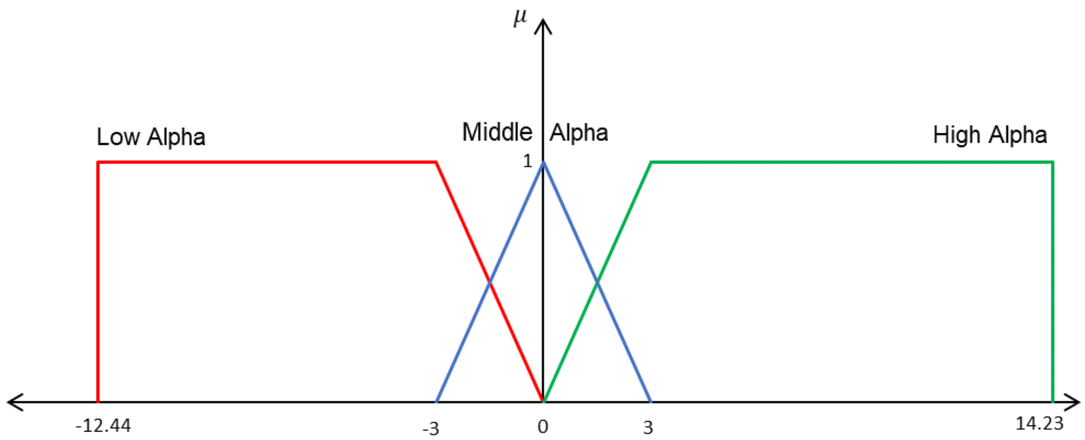


**Figure 1** **Demonstration of fuzzy numbers for linguistic variables Low Alpha, Middle Alpha and High Alpha**
In the figure, $\mu$ is the membership function of a variable on interval [1,0].

For alpha, the following fuzzy numbers are computed: Low_Alpha [-12.44, -12.44, -3, 0]; Middle_Alpha [-3, 0, 0, 3]; and High_Alpha [0, 3, 14.23, 14.23]. As the developments by Stoklasa et al. (2017) of fsQCA and the evidence against the rules (that is, $A \Rightarrow notB$) are also employed, notLow_Alpha is determined as [-3, 0, 14.23, 14.23] and notHigh_Alpha is [-12.44, -12.44, 0, 3]. Linguistic variables Low Alpha, Middle Alpha and High Alpha are presented in Figure 1.

In general, the higher the Sharpe ratio is, the better. However, given no other information of the fund, it cannot be estimated whether a ratio is good or not. The Sharpe ratio should always be compared with another fund or a group of funds. (Morningstar 2015) Thus, these values are chosen arbitrarily using quartiles of the sample's Sharpe ratios to compare the ratios more truthfully. The Sharpe ratios vary from -0.19 to 1.58 in the examined sample. The median ratio for the sample is 0.39 and the average is 0.44. Sharing the values in quartiles, the last value of each quartile are as follows; 0.25, 0.39, 0.63 and 1.58. Here, the following fuzzy numbers are used: Low_Sharpe [-0.19, -0.19, 0.25, 0.39]; Middle_Sharpe [0.25, 0.39, 0.39, 0.63]; High_Sharpe [0.39, 0.63, 1.58, 1.58]. Respectively, notLow_Sharpe is [0.25, 0.39, 1.58, 1.58] and notHigh_Sharpe is [-0.19, -0.19, 0.39, 0.63]. The fuzzy numbers for Low sharpe, Middle sharpe and High sharpe are demonstrated in Figure 2.

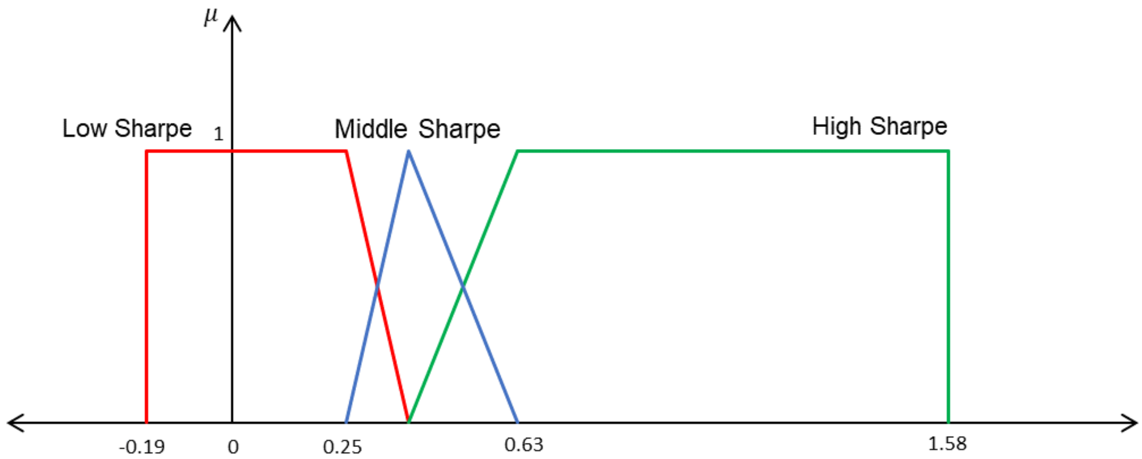


**Figure 2 Demonstration of fuzzy numbers for linguistic variables Low Sharpe, Middle Sharpe and High Sharpe**

In the figure, $\mu$ is the membership function of a variable on interval [1,0].

Next, fuzzy numbers for input variables fund size, manager tenure and the Morningstar Sustainability rating are formed. In past literature, there does not seem to be any general borderlines for large fund size or long manager tenure. Therefore, some researchers (e.g., Kleiman & Jun 1988) divide the sample into quartiles in their evaluations. Thus, the fuzzy numbers are formed by quartiles also for fund size and manager tenure. Summary statistics including of each quartile were presented in Table 5. In the table, the maximum values of each quartile represent the last value of each quartile. The median fund size measured with net assets under management is 235.03 million euros and the average fund size is 564.78 million euros in the sample. Fund sizes in the sample vary from 1.08 million euros to 7124.65 million euros under management. In Figure 3, the following formed fuzzy numbers for fund size are presented: Small_size [1.08, 1.08, 78.39, 235.03]; Middle_size [78.39, 235.03, 235.03, 664.91] and Large_size [235.03, 664.91, 7124.65, 7124.65].

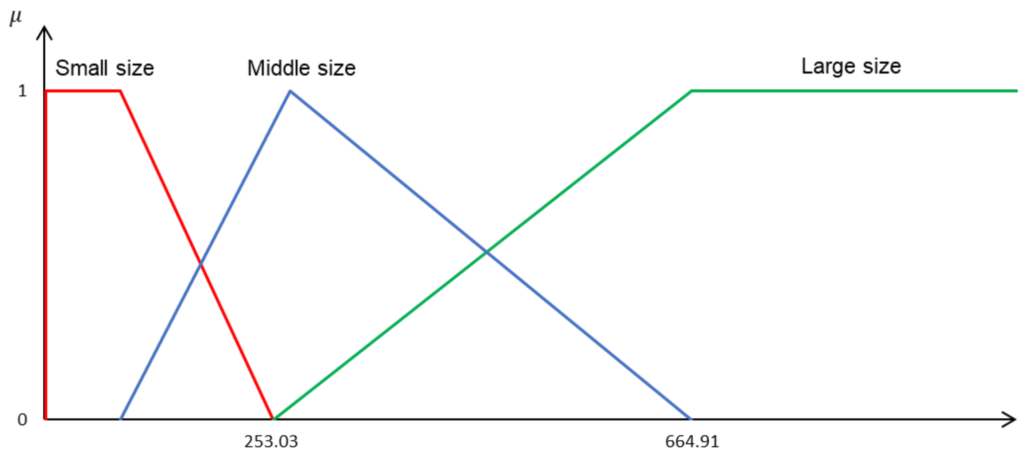


**Figure 3** **Demonstration of fuzzy numbers for linguistic variables Small size, Middle size and Large size**
In the figure, $\mu$ is the membership function of a variable on interval [1,0].

For manager tenure, the median is 7.83 years, and the average is 8.42 years. Manager tenure varies from 0 to 23.58 years. Fuzzy numbers for manager tenure are computed as follows: Short_tenure [0, 0, 3.58, 7.83]; Middle_tenure [3.58, 7.83, 7.83, 12.08] and Long_tenure [7.83, 12.08, 23.58, 23.58]. These fuzzy numbers for linguistic variables Short tenure, Middle tenure and Long tenure are shown in Figure 4.

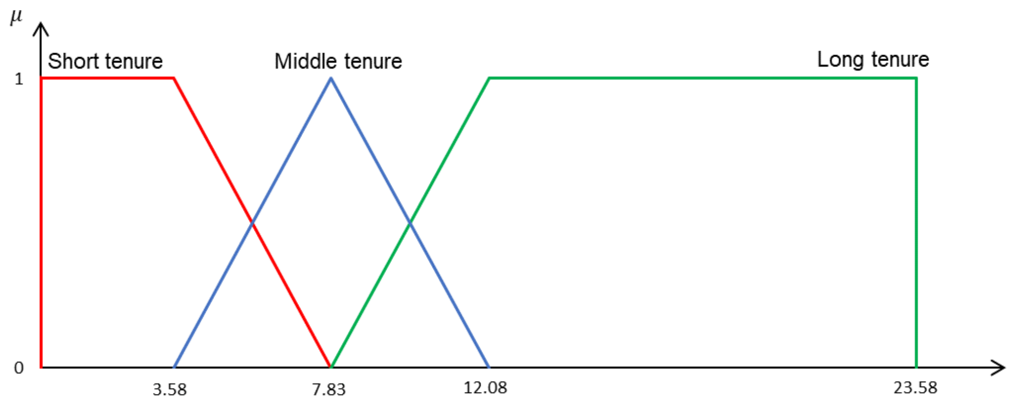


**Figure 4** **Demonstration of fuzzy numbers for linguistic variables Short tenure, Middle tenure and Long tenure**
In the figure, $\mu$ is the membership function of a variable on interval [1,0].

The Morningstar Sustainability rating has values from one to five, with one referring to the lowest level of sustainability and five to the highest level of sustainability. The rating measures a fund's adherence to ESG factors and classifies each fund to a category between one globe (low sustainability) and five globes (high sustainability). The top 10 percent of funds conforming to ESG criteria receive a Sustainability Rating of five, while the bottom 10 percent are categorized with a rating of one. (Morningstar 2019) For sustainability, fuzzy numbers are not formed by quartiles as there already is a scale (1-5) describing the level of a fund’s sustainability. Instead, the following fuzzy numbers are defined: Low_Sustainability [1, 1, 2, 4] and High_Sustainability [2, 4, 5, 5]. These fuzzy numbers are shown in Figure 5.

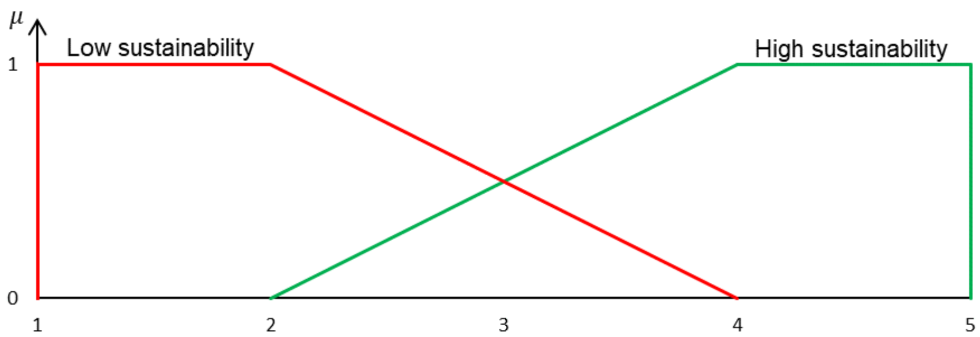


**Figure 5** **Demonstration of fuzzy numbers for linguistic variables Low sustainability and High sustainability**
In the figure, $\mu$ is the membership function of a variable on interval [1,0].

### 4.3.2 Hypothesis demonstration and linguistic rules

As shown in Table 7, we can form a total of nine hypotheses for each variable fund size, manager tenure, and sustainability with respect to the linguistic variables we created. However, as we would in addition have two variables for risk-adjusted returns, which would add the number of hypotheses to 54.

**Table 7 The hypothesis demonstrations with linguistic variables for fund size, manager tenure, and sustainability with risk-adjusted returns**

| | | *Risk-adjusted returns* | | |
|---|---|---|---|---|
| | | low | middle | high |
| | small | $H_{1.1}$ | $H_{1.2}$ | $H_{1.3}$ |
| *Fund size* | middle | $H_{1.4}$ | $H_{1.5}$ | $H_{1.6}$ |
| | large | $H_{1.7}$ | $H_{1.8}$ | $H_{1.9}$ |

| | | *Risk-adjusted returns* | | |
|---|---|---|---|---|
| | | low | middle | high |
| *Manager* | short | $H_{2.1}$ | $H_{2.2}$ | $H_{2.3}$ |
| *tenure* | middle | $H_{2.4}$ | $H_{2.5}$ | $H_{2.6}$ |
| | long | $H_{2.7}$ | $H_{2.8}$ | $H_{2.9}$ |

| | | *Risk-adjusted returns* | | |
|---|---|---|---|---|
| | | low | middle | high |
| | low | $H_{3.1}$ | $H_{3.2}$ | $H_{3.3}$ |
| *Sustainability* | middle | $H_{3.4}$ | $H_{3.5}$ | $H_{3.6}$ |
| | high | $H_{3.7}$ | $H_{3.8}$ | $H_{3.9}$ |

In this chapter, we will focus on examining the hypotheses formed earlier in chapter 2.5. These hypotheses can be read from Table 7 as follows:

$H_{1.7}$: *If fund size is large then risk-adjusted returns are low.*

$H_{2.9}$: *If manager tenure is long then risk-adjusted returns are high.*

$H_{3.9}$: *If Morningstar Sustainability Rating is high then risk-adjusted returns are high.*

Thus, in this analysis, we will focus on the following linguistic rules:

- *Large_size ⇒ Low_sharpe; Large_size ⇒ Low_alpha*
- *Long_tenure ⇒ High_sharpe; Long_tenure ⇒ High_alpha*
- *High_sustainability ⇒ High_sharpe; High_sustainability ⇒ High_alpha*

In addition, to have an inclusive insight of these possible relationships, the same inputs are tested with low (high) risk-adjusted returns as outputs:

- *Large_size ⇒ High_sharpe; Large_size ⇒ High_alpha*
- *Long_tenure ⇒ Low_sharpe; Long_tenure ⇒ Low_alpha*
- *High_sustainability ⇒ Low_sharpe; High_sustainability ⇒ Low_alpha*

# 5 RESULTS

In this chapter, results for the focused hypotheses are presented. The effect of fund size measured with net assets under management, manager tenure, and the Morningstar Sustainability Rating on fund performance measured with risk-adjusted returns Jensen’s Alpha and Sharpe Ratio are investigated.

In the results, consistencies and coverages for F1, F2, F3, and F4 are presented. In the tables, $A \Rightarrow B$ are values in favor of the tested rule and $A \Rightarrow notB$ are values against the tested rules. For F4, a consistency over 0.5 is considered to support in favor of the examined relationship, while in F1, F2, and F3, a high consistency denotes the support in favor of the relationship. In addition, the degree of support and degree of disproof is presented. If $\alpha$ − SUP = 1, there is support in favor of the examined rule to some nonzero degree in the data. The higher the degree for SUP1(A⇒B) is, the more support in favor of the rule there is. Respectively, if $\alpha$− DISP equals to one, there is evidence disproofing the rule in the data. A lower degree of α-plausible support means there is evidence in the data that $B$ is fulfilled completely and $A$ is fulled to some nonzero level.

## 5.1 Large fund size and fund performance

First, the relationship between large fund size and risk-adjusted returns was examined. The tested hypothesis was $H_{1.7}$*: If fund size is large then risk-adjusted returns are low,* which was backed by a significant amount of the past literature on the relationship between fund size and fund performance. The findings from the examination of rules *Large_size ⇒ Low_Sharpe* and *Large_size ⇒ Low_Alpha* are concluded in Table 8.

For both of these relationships, it seems like there is much more support against $A \Rightarrow B$ than in favor of it. Overall, for both low Sharpe Ratios and Alphas, the consistencies and coverages for $A \Rightarrow notB$ are higher. The F3 consistency, which also reflects pure counterevidence, is zero for $A \Rightarrow B$ while for $A \Rightarrow notB$ the consistency is over 0.4. F1 consistency for $A \Rightarrow notB$ is relatively high (0.717 for *not Low_sharpe,* and 0.747 for *not Low_Alpha)* and the F4 consistencies are over 0.7 with decent coverages for both as well. Thus, it seems like the large size could create not-low Sharpe Ratio's or Alphas. $\alpha$ − SUP and $\alpha$− DISP is equal to one in both rules which means there is evidence for both support and disproof for these rules in the sample. The values for SUP1(A⇒B) are quite low (0.191 and 0.131) in both rules compared to the values of DISP1(A⇒B) (0.565 and 0.552), which could imply there is more disproof than support for large size leading to low risk-adjusted returns. However, as $\alpha$ − SUP and $\alpha$− DISP equals to one in these rules, the results might not be that reliable.

**Table 8 Results of the evaluation of the validity of rules Large_size ⇒ Low_Sharpe and Large_size ⇒ Low_Alpha**

| A: Large size ; B: Low sharpe | | | | A: Large size ; B: Low Alpha | | | |
|---|---|---|---|---|---|---|---|
| A ⇒ B | | A ⇒ not B | | A ⇒ B | | A ⇒ not B | |
| F1 consistency = | 0.316 | F1 consistency = | 0.717 | F1 consistency = | 0.290 | F1 consistency = | 0.747 |
| F1 coverage = | 0.308 | F1 coverage = | 0.395 | F1 coverage = | 0.300 | F1 coverage = | 0.399 |
| F2 consistency = | 0.229 | F2 consistency = | 0.631 | F2 consistency = | 0.195 | F2 consistency = | 0.651 |
| F2 coverage = | 0.241 | F2 coverage = | 0.326 | F2 coverage = | 0.227 | F2 coverage = | 0.331 |
| F3 consistency = | 0 | F3 consistency = | 0.401 | F3 consistency = | 0 | F3 consistency = | 0.456 |
| F3 coverage = | 0 | F3 coverage = | 0 | F3 coverage = | 0 | F3 coverage = | 0 |
| F4 consistency = | 0.299 | F4 consistency = | 0.701 | F4 consistency = | 0.272 | F4 consistency = | 0.728 |
| F4 coverage = | 0.292 | F4 coverage = | 0.386 | F4 coverage = | 0.281 | F4 coverage = | 0.389 |
| SUP1(A⇒B) = | 0.191 | DISP1(A⇒B) = | 0.565 | SUP1(A⇒B) = | 0.131 | DISP1(A⇒B) = | 0.552 |
| SUP0.9(A⇒B) = | 0.202 | DISP0.9(A⇒B) = | 0.577 | SUP0.9(A⇒B) = | 0.148 | DISP0.9(A⇒B) = | 0.592 |
| SUP0.8(A⇒B) = | 0.209 | DISP0.8(A⇒B) = | 0.609 | SUP0.8(A⇒B) = | 0.169 | DISP0.8(A⇒B) = | 0.640 |
| SUP0.7(A⇒B) = | 0.234 | DISP0.7(A⇒B) = | 0.643 | SUP0.7(A⇒B) = | 0.206 | DISP0.7(A⇒B) = | 0.681 |
| SUP0.6(A⇒B) = | 0.256 | DISP0.6(A⇒B) = | 0.682 | SUP0.6(A⇒B) = | 0.243 | DISP0.6(A⇒B) = | 0.704 |
| SUP0.5(A⇒B) = | 0.294 | DISP0.5(A⇒B) = | 0.715 | SUP0.5(A⇒B) = | 0.280 | DISP0.5(A⇒B) = | 0.720 |
| SUP0.4(A⇒B) = | 0.318 | DISP0.4(A⇒B) = | 0.744 | SUP0.4(A⇒B) = | 0.296 | DISP0.4(A⇒B) = | 0.764 |
| SUP0.3(A⇒B) = | 0.357 | DISP0.3(A⇒B) = | 0.766 | SUP0.3(A⇒B) = | 0.319 | DISP0.3(A⇒B) = | 0.794 |
| SUP0.2(A⇒B) = | 0.391 | DISP0.2(A⇒B) = | 0.791 | SUP0.2(A⇒B) = | 0.360 | DISP0.2(A⇒B) = | 0.831 |
| SUP0.1(A⇒B) = | 0.423 | DISP0.1(A⇒B) = | 0.798 | SUP0.1(A⇒B) = | 0.408 | DISP0.1(A⇒B) = | 0.852 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 | SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 | α-SUP = | 1 | α-DISP = | 1 |

In addition, the same input *Large_size* for outputs *High_Sharpe* and *High_Alpha* is tested. As the previous results suggested that large fund size could create not-low risk-adjusted returns, one could wonder if large fund size would then lead to high risk-adjusted returns. The results of the examination of these rules are summarized in Table 9. Although there does not seem to be much support for *Large_size ⇒ High_sharpe,* there is some weak evidence (e.g., F1 cons. 0.438; F2 cons. 0.358; F4 cons. 0.425) in favor of the rule. In addition, there is more evidence in favor of *Large_size ⇒ High_sharpe* than there was for *Large_size ⇒ Low_Sharpe* in Table 8. However, none of the consistencies are at a sufficient level for a strong relationship, and F3 consistency, which also reflects pure counterevidence, is zero. F4 consistency for $A \Rightarrow notB$ is over 0.5 (0.575) with a considerably high coverage. In addition, α − SUP and α – DISP are both one and the degree of disproof is higher than the degree of support.

**Table 9 Results of the evaluation of the validity of rules Large_size ⇒ High_Sharpe and Large_size ⇒ High_Alpha**

| A: Large size ; B: High sharpe | | | | A: Large size ; B: High Alpha | | | |
|---|---|---|---|---|---|---|---|
| A ⇒ B | | A ⇒ not B | | A ⇒ B | | A ⇒ not B | |
| F1 consistency = | 0.438 | F1 consistency = | 0.588 | F1 consistency = | 0.318 | F1 consistency = | 0.708 |
| F1 coverage = | 0.423 | F1 coverage = | 0.326 | F1 coverage = | 0.392 | F1 coverage = | 0.349 |
| F2 consistency = | 0.358 | F2 consistency = | 0.509 | F2 consistency = | 0.236 | F2 consistency = | 0.625 |
| F2 coverage = | 0.351 | F2 coverage = | 0.263 | F2 coverage = | 0.314 | F2 coverage = | 0.288 |
| F3 consistency = | 0 | F3 consistency = | 0.151 | F3 consistency = | 0 | F3 consistency = | 0.389 |
| F3 coverage = | 0 | F3 coverage = | 0 | F3 coverage = | 0 | F3 coverage = | 0 |
| F4 consistency = | 0.425 | F4 consistency = | 0.575 | F4 consistency = | 0.305 | F4 consistency = | 0.695 |
| F4 coverage = | 0.410 | F4 coverage = | 0.318 | F4 coverage = | 0.376 | F4 coverage = | 0.342 |
| SUP 1(A⇒B) = | 0.302 | DISP1(A⇒B) = | 0.448 | SUP1(A⇒B) = | 0.185 | DISP1(A⇒B) = | 0.448 |
| SUP0.9(A⇒B) = | 0.322 | DISP0.9(A⇒B) = | 0.473 | SUP0.9(A⇒B) = | 0.212 | DISP0.9(A⇒B) = | 0.531 |
| SUP0.8(A⇒B) = | 0.343 | DISP0.8(A⇒B) = | 0.483 | SUP0.8(A⇒B) = | 0.222 | DISP0.8(A⇒B) = | 0.606 |
| SUP0.7(A⇒B) = | 0.371 | DISP0.7(A⇒B) = | 0.528 | SUP0.7(A⇒B) = | 0.249 | DISP0.7(A⇒B) = | 0.674 |
| SUP0.6(A⇒B) = | 0.395 | DISP0.6(A⇒B) = | 0.558 | SUP0.6(A⇒B) = | 0.272 | DISP0.6(A⇒B) = | 0.708 |
| SUP0.5(A⇒B) = | 0.431 | DISP0.5(A⇒B) = | 0.581 | SUP0.5(A⇒B) = | 0.279 | DISP0.5(A⇒B) = | 0.721 |
| SUP0.4(A⇒B) = | 0.442 | DISP0.4(A⇒B) = | 0.605 | SUP0.4(A⇒B) = | 0.292 | DISP0.4(A⇒B) = | 0.728 |
| SUP0.3(A⇒B) = | 0.472 | DISP0.3(A⇒B) = | 0.629 | SUP0.3(A⇒B) = | 0.333 | DISP0.3(A⇒B) = | 0.751 |
| SUP0.2(A⇒B) = | 0.517 | DISP0.2(A⇒B) = | 0.657 | SUP0.2(A⇒B) = | 0.394 | DISP0.2(A⇒B) = | 0.778 |
| SUP0.1(A⇒B) = | 0.527 | DISP0.1(A⇒B) = | 0.678 | SUP0.1(A⇒B) = | 0.469 | DISP0.1(A⇒B) = | 0.788 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 | SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 | α-SUP = | 1 | α-DISP = | 1 |

Analyzing the results for *Large_size ⇒ High_Alpha*, evidence against the rule is stronger with an F1 value of 0.708 and F4 value of 0.695. As α − SUP and α − DISP equal to one in both rules, these results may not be reliable. Additionally, as SUP1(A⇒B) is only 0.185 and DISP1(A⇒B) is 0.448, there is more evidence disproofing the relation than supporting it. To conclude, based on these results, there is no strong proof of a relationship between large size leading to low risk-adjusted returns. There does seem to be a little more support for large size leading to high risk-adjusted returns, but as there was more evidence against the rules Large_size ⇒ High_Sharpe and Large_size ⇒ High_Alpha and no strong proof in favor of them, any conclusions cannot be drawn.

### *5.2 Long manager tenure and fund performance*

Next, the hypothesis $H_{2.9}$: *If manager tenure is long then risk-adjusted returns are high* was examined. The results are concluded in Table 10. For both tested rules, there is more support against them, and the values for $A \Rightarrow not\ B$ are higher compared to the values for $A \Rightarrow B$. With an F1 consistency of 0.746 and an F4 consistency of 0.729, there seems to be rather strong evidence for *Long_tenure ⇒notHigh_Alpha*. Similarly, for *Long_tenure ⇒ notHigh_Sharpe,* the F4 consistency is 0.658 and DISP1(A⇒B) is 0.513 while SUP1(A⇒B) is only 0.216. Based on these findings, there does not seem

**Table 10 Results of the evaluation of the validity of rules Long_tenure ⇒ High_Sharpe and Long_tenure ⇒ High_Alpha**

| A: Long tenure; B: High sharpe | | | | A: Long tenure; B: High Alpha | | | |
|---|---|---|---|---|---|---|---|
| | A ⇒ B | | A ⇒ not B | | A ⇒ B | | A ⇒ not B |
| F1 consistency = | 0.355 | F1 consistency = | 0.670 | F1 consistency = | 0.287 | F1 consistency = | 0.746 |
| F1 coverage = | 0.367 | F1 coverage = | 0.396 | F1 coverage = | 0.377 | F1 coverage = | 0.393 |
| F2 consistency = | 0.278 | F2 consistency = | 0.593 | F2 consistency = | 0.193 | F2 consistency = | 0.652 |
| F2 coverage = | 0.291 | F2 coverage = | 0.330 | F2 coverage = | 0.283 | F2 coverage = | 0.329 |
| F3 consistency = | 0 | F3 consistency = | 0.315 | F3 consistency = | 0 | F3 consistency = | 0.459 |
| F3 coverage = | 0 | F3 coverage = | 0 | F3 coverage = | 0 | F3 coverage = | 0 |
| F4 consistency = | 0.342 | F4 consistency = | 0.658 | F4 consistency = | 0.271 | F4 consistency = | 0.729 |
| F4 coverage = | 0.353 | F4 coverage = | 0.389 | F4 coverage = | 0.355 | F4 coverage = | 0.384 |
| SUP1(A⇒B) = | 0.216 | DISP1(A⇒B) = | 0.513 | SUP1(A⇒B) = | 0.132 | DISP1(A⇒B) = | 0.501 |
| SUP0.9(A⇒B) = | 0.237 | DISP0.9(A⇒B) = | 0.538 | SUP0.9(A⇒B) = | 0.166 | DISP0.9(A⇒B) = | 0.567 |
| SUP0.8(A⇒B) = | 0.251 | DISP0.8(A⇒B) = | 0.571 | SUP0.8(A⇒B) = | 0.178 | DISP0.8(A⇒B) = | 0.627 |
| SUP0.7(A⇒B) = | 0.298 | DISP0.7(A⇒B) = | 0.613 | SUP0.7(A⇒B) = | 0.197 | DISP0.7(A⇒B) = | 0.687 |
| SUP0.6(A⇒B) = | 0.319 | DISP0.6(A⇒B) = | 0.637 | SUP0.6(A⇒B) = | 0.235 | DISP0.6(A⇒B) = | 0.726 |
| SUP0.5(A⇒B) = | 0.340 | DISP0.5(A⇒B) = | 0.662 | SUP0.5(A⇒B) = | 0.246 | DISP0.5(A⇒B) = | 0.754 |
| SUP0.4(A⇒B) = | 0.363 | DISP0.4(A⇒B) = | 0.681 | SUP0.4(A⇒B) = | 0.274 | DISP0.4(A⇒B) = | 0.765 |
| SUP0.3(A⇒B) = | 0.387 | DISP0.3(A⇒B) = | 0.702 | SUP0.3(A⇒B) = | 0.320 | DISP0.3(A⇒B) = | 0.803 |
| SUP0.2(A⇒B) = | 0.429 | DISP0.2(A⇒B) = | 0.749 | SUP0.2(A⇒B) = | 0.373 | DISP0.2(A⇒B) = | 0.822 |
| SUP0.1(A⇒B) = | 0.462 | DISP0.1(A⇒B) = | 0.763 | SUP0.1(A⇒B) = | 0.433 | DISP0.1(A⇒B) = | 0.841 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 | SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 | α-SUP = | 1 | α-DISP = | 1 |

to be much support for the relationship between long manager tenure and high risk-adjusted returns. On the contrary, these results are leaning more towards supporting the relation between long manager tenure and not-high risk-adjusted returns. Thus, the relationship between long tenure and low risk-adjusted returns is also tested. Here, rules *Long_tenure ⇒ Low_Sharpe* and *Long_tenure ⇒ Low_Alpha* were accessed. The results of these findings are concluded in Table 11.

**Table 11 Results of the evaluation of the validity of rules Long_tenure ⇒ Low_Sharpe and Long_tenure ⇒ Low_Alpha**

| A: Long tenure; B: Low sharpe | | | |
|---|---|---|---|
| A ⇒ B | | A ⇒ not B | |
| F1 consistency = | 0.391 | F1 consistency = | 0.644 |
| F1 coverage = | 0.407 | F1 coverage = | 0.379 |
| F2 consistency = | 0.305 | F2 consistency = | 0.558 |
| F2 coverage = | 0.336 | F2 coverage = | 0.305 |
| F3 consistency = | 0 | F3 consistency = | 0.253 |
| F3 coverage = | 0 | F3 coverage = | 0 |
| F4 consistency = | 0.374 | F4 consistency = | 0.626 |
| F4 coverage = | 0.389 | F4 coverage = | 0.368 |
| SUP1(A⇒B) = | 0.254 | DISP1(A⇒B) = | 0.505 |
| SUP0.9(A⇒B) = | 0.266 | DISP0.9(A⇒B) = | 0.509 |
| SUP0.8(A⇒B) = | 0.287 | DISP0.8(A⇒B) = | 0.535 |
| SUP0.7(A⇒B) = | 0.326 | DISP0.7(A⇒B) = | 0.574 |
| SUP0.6(A⇒B) = | 0.340 | DISP0.6(A⇒B) = | 0.594 |
| SUP0.5(A⇒B) = | 0.377 | DISP0.5(A⇒B) = | 0.642 |
| SUP0.4(A⇒B) = | 0.406 | DISP0.4(A⇒B) = | 0.660 |
| SUP0.3(A⇒B) = | 0.426 | DISP0.3(A⇒B) = | 0.674 |
| SUP0.2(A⇒B) = | 0.465 | DISP0.2(A⇒B) = | 0.713 |
| SUP0.1(A⇒B) = | 0.491 | DISP0.1(A⇒B) = | 0.734 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 |

| A: Long tenure; B: Low Alpha | | | |
|---|---|---|---|
| A ⇒ B | | A ⇒ not B | |
| F1 consistency = | 0.346 | F1 consistency = | 0.680 |
| F1 coverage = | 0.382 | F1 coverage = | 0.387 |
| F2 consistency = | 0.252 | F2 consistency = | 0.585 |
| F2 coverage = | 0.310 | F2 coverage = | 0.319 |
| F3 consistency = | 0 | F3 consistency = | 0.333 |
| F3 coverage = | 0 | F3 coverage = | 0 |
| F4 consistency = | 0.333 | F4 consistency = | 0.667 |
| F4 coverage = | 0.368 | F4 coverage = | 0.380 |
| SUP1(A⇒B) = | 0.189 | DISP1(A⇒B) = | 0.499 |
| SUP0.9(A⇒B) = | 0.207 | DISP0.9(A⇒B) = | 0.534 |
| SUP0.8(A⇒B) = | 0.233 | DISP0.8(A⇒B) = | 0.554 |
| SUP0.7(A⇒B) = | 0.265 | DISP0.7(A⇒B) = | 0.608 |
| SUP0.6(A⇒B) = | 0.296 | DISP0.6(A⇒B) = | 0.658 |
| SUP0.5(A⇒B) = | 0.321 | DISP0.5(A⇒B) = | 0.679 |
| SUP0.4(A⇒B) = | 0.342 | DISP0.4(A⇒B) = | 0.710 |
| SUP0.3(A⇒B) = | 0.398 | DISP0.3(A⇒B) = | 0.735 |
| SUP0.2(A⇒B) = | 0.446 | DISP0.2(A⇒B) = | 0.767 |
| SUP0.1(A⇒B) = | 0.466 | DISP0.1(A⇒B) = | 0.793 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 |

Again, there is more support against these rules than in favor of them. Although according to the results in Table 10 there is more support for long-tenured managers creating not-high risk-adjusted returns, results in Table 11 show that there is also evidence for a long tenure of a manager creating not-low risk-adjusted returns. As F4 consistencies are over 0.5 against both rules and the overall consistencies are stronger against both rules, long tenure of a manager could prevent from having low risk-adjusted returns. Because results shown in Tables 10 and 11 are both showing more disproof for the rules and α − SUP and α – DISP equal to one in all tested rules, these results do not seem trustworthy.

## *5.3 High sustainability and fund performance*

The hypothesis for the possible relationship between Morningstar Sustainability Rating and fund performance $H_{3.9}$*: If Morningstar Sustainability Rating is high then risk-adjusted returns are high* was based on findings from previous literature. The results for the rules *High_Sustainability ⇒ High_Sharpe* and *High_sustainability ⇒ High_Alpha* are summarized in Table 12.

**Table 12 Results of the evaluation of the validity of rules High_sustainability ⇒ High_Sharpe and High_sustainability ⇒ High_Alpha**

| A: High_sustainability; B: High_sharpe | | | | A: High_sustainability; B: High_Alpha | | | |
|---|---|---|---|---|---|---|---|
| A ⇒ B | | A ⇒ not B | | A ⇒ B | | A ⇒ not B | |
| F1 consistency = | 0.473 | F1 consistency = | 0.587 | F1 consistency = | 0.357 | F1 consistency = | 0.698 |
| F1 coverage = | 0.718 | F1 coverage = | 0.510 | F1 coverage = | 0.690 | F1 coverage = | 0.541 |
| F2 consistency = | 0.378 | F2 consistency = | 0.493 | F2 consistency = | 0.263 | F2 consistency = | 0.604 |
| F2 coverage = | 0.498 | F2 coverage = | 0.265 | F2 coverage = | 0.464 | F2 coverage = | 0.304 |
| F3 consistency = | 0 | F3 consistency = | 0.114 | F3 consistency = | 0 | F3 consistency = | 0.341 |
| F3 coverage = | 0.344 | F3 coverage = | 0 | F3 coverage = | 0.272 | F3 coverage = | 0.038 |
| F4 consistency = | 0.443 | F4 consistency = | 0.557 | F4 consistency = | 0.329 | F4 consistency = | 0.671 |
| F4 coverage = | 0.672 | F4 coverage = | 0.484 | F4 coverage = | 0.636 | F4 coverage = | 0.519 |
| SUP1(A⇒B) = | 0.323 | DISP1(A⇒B) = | 0.441 | SUP1(A⇒B) = | 0.203 | DISP1(A⇒B) = | 0.483 |
| SUP0.9(A⇒B) = | 0.340 | DISP0.9(A⇒B) = | 0.460 | SUP0.9(A⇒B) = | 0.230 | DISP0.9(A⇒B) = | 0.530 |
| SUP0.8(A⇒B) = | 0.357 | DISP0.8(A⇒B) = | 0.481 | SUP0.8(A⇒B) = | 0.247 | DISP0.8(A⇒B) = | 0.589 |
| SUP0.7(A⇒B) = | 0.390 | DISP0.7(A⇒B) = | 0.502 | SUP0.7(A⇒B) = | 0.270 | DISP0.7(A⇒B) = | 0.633 |
| SUP0.6(A⇒B) = | 0.418 | DISP0.6(A⇒B) = | 0.523 | SUP0.6(A⇒B) = | 0.302 | DISP0.6(A⇒B) = | 0.669 |
| SUP0.5(A⇒B) = | 0.456 | DISP0.5(A⇒B) = | 0.553 | SUP0.5(A⇒B) = | 0.314 | DISP0.5(A⇒B) = | 0.686 |
| SUP0.4(A⇒B) = | 0.477 | DISP0.4(A⇒B) = | 0.582 | SUP0.4(A⇒B) = | 0.331 | DISP0.4(A⇒B) = | 0.698 |
| SUP0.3(A⇒B) = | 0.498 | DISP0.3(A⇒B) = | 0.610 | SUP0.3(A⇒B) = | 0.371 | DISP0.3(A⇒B) = | 0.730 |
| SUP0.2(A⇒B) = | 0.519 | DISP0.2(A⇒B) = | 0.643 | SUP0.2(A⇒B) = | 0.411 | DISP0.2(A⇒B) = | 0.753 |
| SUP0.1(A⇒B) = | 0.540 | DISP0.1(A⇒B) = | 0.660 | SUP0.1(A⇒B) = | 0.470 | DISP0.1(A⇒B) = | 0.774 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 | SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 | α-SUP = | 1 | α-DISP = | 1 |

For the relationship High_sustainability ⇒ High_sharpe, there does not seem to be strong support in the data, and there are no significant differences in consistencies in favor and against the rule. The consistencies for *High_Sustainability ⇒ notHigh_Sharpe* seem to be slightly higher, and F4 consistency is over 0.5, which could indicate high sustainability ratings creating not-high values for the Sharpe Ratio. Likewise, when focusing on the rule *High_sustainability ⇒ High_Alpha*, the consistencies against the rule are a little higher with an F1 consistency of 0.698 and F4 consistency of 0.671. Additionally, F3 consistency is 0 for both *High_Sustainability ⇒ High_Sharpe* and *High_sustainability ⇒ High_Alpha*.

Observing the degrees of support and disproof, both α − SUP and α − DISP equal to one in both rules, which means there is both support and disproof for the rules. Additionally, especially for the rule *High_Sustainability ⇒ High_Sharpe*, the degree of support (0.323) and degree of disproof (0.441) do not differ substantially. In conclusion, there does not seem to be a pattern pointing to funds with high Morningstar Sustainability Ratings having higher Sharpe Ratios or Alphas. The results do not show a strong relationship in favor or against high sustainability leading to high risk-adjusted returns.

To further investigate the relationship between high Morningstar Sustainability Rating and fund performance, we also tested the rules *High_Sustainability ⇒ Low_Sharpe* and *High_sustainability ⇒ Low_Alpha.* The results for these rules are summarized in Table 13. There is not much support in favor of either of these rules; overall consistencies seem low and F3 consistency is 0 for both *High_Sustainability ⇒ Low_Sharpe and High_sustainability ⇒ Low_Alpha*. However, when focusing on proof against the rule $(A \Rightarrow not\ B)$, there are rather strong results. The consistencies are overall higher compared to consistencies in favor of these rules, and there is some coverage to support these consistencies as well. Consistencies for F1 (0.725 and 0.727) and F2 (0.634 and 0.616), are also moderately high. For F4, the consistencies are even considerably high (0.697 and 0.696), with coverages over 0.5. Again, it must be pointed that as α − SUP and α − DISP equal to one in both rules, the relationships might not be reliable. However, SUP1(A⇒B) is only 0.211 for *High_Sustainability ⇒ Low_Sharpe* and 0.173 for *High_sustainability ⇒ Low_Alpha,* while DISP1(A⇒B) is over 0.5 for both rules.

**Table 13 Results of the evaluation of the validity of rules High_sustainability ⇒ Low_Sharpe and High_sustainability ⇒ Low_Alpha**

| A: High_sustainability; B: Low_sharpe | | | | A: High_sustainability; B: Low_Alpha | | | |
|---|---|---|---|---|---|---|---|
| A ⇒ B | | A ⇒ not B | | A ⇒ B | | A ⇒ not B | |
| F1 consistency = | 0.331 | F1 consistency = | 0.725 | F1 consistency = | 0.335 | F1 consistency = | 0.727 |
| F1 coverage = | 0.507 | F1 coverage = | 0.627 | F1 coverage = | 0.543 | F1 coverage = | 0.609 |
| F2 consistency = | 0.240 | F2 consistency = | 0.634 | F2 consistency = | 0.225 | F2 consistency = | 0.616 |
| F2 coverage = | 0.283 | F2 coverage = | 0.387 | F2 coverage = | 0.289 | F2 coverage = | 0.381 |
| F3 consistency = | 0 | F3 consistency = | 0.394 | F3 consistency = | 0 | F3 consistency = | 0.392 |
| F3 coverage = | 0 | F3 coverage = | 0.205 | F3 coverage = | 0 | F3 coverage = | 0.166 |
| F4 consistency = | 0.303 | F4 consistency = | 0.697 | F4 consistency = | 0.304 | F4 consistency = | 0.696 |
| F4 coverage = | 0.464 | F4 coverage = | 0.603 | F4 coverage = | 0.493 | F4 coverage = | 0.583 |
| SUP1(A⇒B) = | 0.211 | DISP1(A⇒B) = | 0.586 | SUP1(A⇒B) = | 0.173 | DISP1(A⇒B) = | 0.521 |
| SUP0.9(A⇒B) = | 0.213 | DISP0.9(A⇒B) = | 0.599 | SUP0.9(A⇒B) = | 0.179 | DISP0.9(A⇒B) = | 0.559 |
| SUP0.8(A⇒B) = | 0.222 | DISP0.8(A⇒B) = | 0.614 | SUP0.8(A⇒B) = | 0.194 | DISP0.8(A⇒B) = | 0.589 |
| SUP0.7(A⇒B) = | 0.249 | DISP0.7(A⇒B) = | 0.639 | SUP0.7(A⇒B) = | 0.234 | DISP0.7(A⇒B) = | 0.639 |
| SUP0.6(A⇒B) = | 0.264 | DISP0.6(A⇒B) = | 0.677 | SUP0.6(A⇒B) = | 0.268 | DISP0.6(A⇒B) = | 0.688 |
| SUP0.5(A⇒B) = | 0.304 | DISP0.5(A⇒B) = | 0.711 | SUP0.5(A⇒B) = | 0.297 | DISP0.5(A⇒B) = | 0.703 |
| SUP0.4(A⇒B) = | 0.323 | DISP0.4(A⇒B) = | 0.736 | SUP0.4(A⇒B) = | 0.312 | DISP0.4(A⇒B) = | 0.732 |
| SUP0.3(A⇒B) = | 0.361 | DISP0.3(A⇒B) = | 0.751 | SUP0.3(A⇒B) = | 0.365 | DISP0.3(A⇒B) = | 0.766 |
| SUP0.2(A⇒B) = | 0.386 | DISP0.2(A⇒B) = | 0.778 | SUP0.2(A⇒B) = | 0.411 | DISP0.2(A⇒B) = | 0.806 |
| SUP0.1(A⇒B) = | 0.401 | DISP0.1(A⇒B) = | 0.787 | SUP0.1(A⇒B) = | 0.441 | DISP0.1(A⇒B) = | 0.821 |
| SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 | SUP0.0(A⇒B) = | 1 | DISP0.0(A⇒B) = | 1 |
| α-SUP = | 1 | α-DISP = | 1 | α-SUP = | 1 | α-DISP = | 1 |

In conclusion, the results for *High_sustainability ⇒ High_Sharpe* and *High_sustainability ⇒ High_Alpha* concluded in Table 12 did not show strong evidence in favor or against the rules. Examining the rules *High_sustainability ⇒ Low_Sharpe* and *High_sustainability ⇒ Low_Alpha*, considerably strong evidence was found against these rules indicating a relationship between high sustainability and not-low risk-adjusted returns. These results suggest that although funds with higher sustainability may not reach higher risk-adjusted returns, they might prevent the fund from having low ones.

# 6 CONCLUSIONS

This thesis investigated if there exists a relationship between mutual fund characteristics manager tenure, fund size and sustainability, and mutual fund performance measured with Jensen's Alpha and Sharpe Ratio. Methodologically, a novel approach to mutual fund performance evaluation was made by using fuzzy set qualitative comparative analysis (fsQCA) by Ragin (2008), with its enhancements developed by Stoklasa et al. (2017, 2018). The objective of this thesis was to examine the possible relationships between the Morningstar Sustainability Rating, fund size, manager tenure, and mutual fund performance. While these relationships have been studied several times, the evidence in the area is inconclusive. A collective literature review presented a range of results and suggestions for the relationships mentioned above, and these specific characteristics were chosen for the evaluation. Mutual fund performance was measured with three-year average Sharpe Ratios and Jensen's Alphas. The examined sample included 429 mutual growth funds registered in Europe and the studied period was from March 2018 to March 2021.

Findings in past research suggested a negative relationship between large fund size and fund performance. The results did not find a strong pattern of large fund sizes creating low risk-adjusted returns. On the contrary, there was no strong evidence of large fund size leading to high risk-adjusted returns either. Results did not support a positive relationship between long manager tenure and fund performance, but they could implicate that the long tenure of a manager could prevent the fund from having low risk-adjusted returns. On the other hand, results showed more evidence for a long tenure of a manager creating not-high performance. The results for either fund size or manager tenure did not show strong proof of relationships.

Results on the effect of high Morningstar Sustainability Ratings did produce interesting findings. There seems to be some support for high sustainability creating high risk-adjusted returns, but the support for high sustainability creating not-high risk-adjusted returns is stronger. However, when examining a rule *if sustainability is high then performance is low*, the evidence against it is stronger. Based on the results, there seems to be a considerably strong relationship between high Morningstar Sustainability rating and not-low risk-adjusted returns implicating that the high sustainability of a fund could avoid it from having low risk-adjusted returns. In other words, investing in funds with high Morningstar Sustainability ratings may not create abnormal financial performance, but it can help avoid poor financial performance.

The current COVID19 pandemic could partially emphasize these results as it has caused the studied period to be economically unstable. Early in the pandemic, sustainable market actors Bloomberg, Morningstar and MSCI reported ESG funds and indices outperforming conventional funds as they were losing less value than their conventional indices (Boffo and Patalano (2020). In support of this, Nofsinger and Varma (2014) found socially responsible funds outperforming during times of market crisis and underperforming during times of non-crisis. Our results outline these observations since high sustainability helped funds avoid low risk-adjusted returns during the studied period. However, as this study examined only three-year values of the whole period, and thus did not consider economically more stable periods, these results do not offer proof of the results being affected by recent market turmoil.

The limitations of this study occurred mostly from the survivorship bias in the data sample. As Morningstar Mutual Fund Screener does not hold data for dead funds, all funds that did not survive over the period (March 2018 – March 2021) were eliminated. Furthermore, look-ahead bias may have occurred in the sample since values for fund size (net assets), manager tenure, and the Morningstar Sustainability Rating were current values, while risk-adjusted returns were three-year averages. Also, the computed fuzzy numbers were mostly created arbitrarily and not based on theoretical information, which can weaken the results. This thesis did not study other mutual fund types than equity funds due to the low amount of available data for funds from other categories, such as bond or income funds.

Future research could extend these findings by studying the role of sustainability during periods of market crisis and periods of non-crisis. By doing this, the overall impact of sustainable investing on financial performance could be observed. In addition, a bias-free data sample should be used to validate these findings. The lack of long-term data for the Morningstar Sustainability Ratings published in 2016 propose avenues for future research to explore the effects of the rating on long-term mutual fund performance. Methodologically, there is much room for applying fsQCA and its enhancements developed by Stoklasa et al. (2017, 2018) in future mutual fund research and performance evaluation. Based on this study, the methodology works well for mutual fund performance research and succeeded in identifying patterns in the sample.

## References


Barillas, F. and Shanken, J. (2017), "Which Alpha?", *The Review of Financial Studies,* Vol. 30, No. 4, pp. 1316-1338.

Barnett, M.L. and Salomon, R.M. (2006), "Beyond Dichotomy: The Curvilinear Relationship between Social Responsibility and Financial Performance", *Strategic Management Journal*, Vol. 27, No. 11, pp. 1101-1122.

Basile, I. and Ferrari, P. (2016), Asset Management and Institutional Investors. *Springer International Publishing*, Cham.

Beckers, S.E. and Vaughan, G. (2001), "Small Is Beautiful"*, Journal of Portfolio Management,* Vol. 27, No. 4, pp. 9-17.

Bello, Z.Y. (2005), "Socially Responsible Investing and Portfolio Diversification", *The Journal of Financial Research*, Vol. 28, No. 1, pp. 41-57.

Boffo, R. and Patalano R. (2020), “ESG Investing: Practices, Progress and Challenges”, OECD Paris. Available: [http://www.oecd.org/finance/ESG-Investing-Practices-Progress-Challenges.pdf](http://www.oecd.org/finance/ESG-Investing-Practices-Progress-Challenges.pdf)

Brooks, M. and Tompkins, D.L. (2002), "Mutual funds' risk adjusted performance", *Academy of banking studies journal*, Vol. 1, pp. 111.

Brown, S.J., Goetzmann, W., Ibbotson, R.G. and Ross, S.A. (1992), "Survivorship Bias in Performance Studies", *The Review of financial studies*, Vol. 5, No. 4, pp. 553-580.

Carhart, M.M. (1997), "On Persistence in Mutual Fund Performance", *The Journal of finance (New York)*, Vol. 52, No. 1, pp. 57-82.

Chan, H.W.H., Faff, R.W., Gallagher, D.R. and Looi, A. (2009), "Fund Size, Transaction Costs and Performance: Size Matters", *Australian Journal of Management*, Vol. 34, No. 1, pp. 73-96.

Chen, J., Hong, H., Huang, M. and Kubik, J.D. (2004), "Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization"*, The American Economic Review,* Vol. 94, No. 5, pp. 1276-1302.

Costa, B.A., Jakob, K. and Porter, G.E. (2006), "Mutual Fund Performance and Changing Market Trends 1990-2001: Does Manager Experience Matter?", *The Journal of investing*, Vol. 15, No. 2, pp. 79-86.

Cuthbertson, K., Nitzsche, D. and O'Sullivan, N. (2016), "A review of behavioural and management effects in mutual fund performance*", International Review of Financial Analysis*, Vol. 44, pp. 162-176.

Dolvin, S., Fulkerson, J. and Krukover, A. (2019), "Do “Good Guys” Finish Last? The Relationship between Morningstar Sustainability Ratings and Mutual Fund Performance"*, The Journal of Investing,* Vol. 28, No. 2, pp. 77-91.

Ejara, D.D. and Nag, R. (2011), "Managerial Tenure And Mutual Fund Performance: Evidence From Index Funds"*, Journal of Business & Economics Research (Littleton, Colo.),* Vol. 7, No. 5.

Fama, E.F. and French, K.R. (2004), "The Capital Asset Pricing Model: Theory and Evidence", *The Journal of Economic Perspectives*, Vol. 18, No. 3, pp. 25-46.

Ferreira, M.A., Keswani, A., Miguel, A.F. and Ramos, S.B. (2013), "The Determinants of Mutual Fund Performance: A Cross-Country Study", *Review of Finance*, Vol. 17, No. 2, pp. 483-525.

Filbeck, G. and Tompkins, D.L. (2004), "Management Tenure and Risk-Adjusted Performance of Mutual Funds", *The Journal of Investing*, Vol. 13, No. 2, pp. 72-80.

Fortin, R., Michelson, S. and Jordan-Wagner, J. (1999), "Does mutual fund manager tenure matter?", *Journal of Financial Planning*, Vol. 12, No. 7, pp. 72.

GSIA (2018), “2018 Global Sustainable Investment Review” GLOBAL SUSTAINABLE INVESTMENT ALLIANCE. [Web document] [Accessed 31.3.2021]. Available: http://www.gsi-alliance.org/wp-content/uploads/2019/03/GSIR_Review2018.3.28.pdf

Golec, J. (1996), "The effects of mutual fund managers' characteristics on their portfolio performance, risk and fees", *Financial Services Review* (Greenwich, Conn.), Vol. 5, No. 2, pp. 133-147.

Graham, J.E., Lassala, C. and Ribeiro-Navarrete, B. (2019), "A fuzzy-set analysis of conditions influencing mutual fund performance", *International Review of Economics & Finance*, Vol. 61, pp. 324-336.

Graham, J.E., Lassala, C. and Ribeiro Navarrete, B. (2020), "Influences on mutual fund performance: comparing US and Europe using qualitative comparative analysis", *Ekonomska istraživanja*, Vol. 33, No. 1, pp. 3049-3070.

Hamilton, S., Jo, H. and Statman, M. (1993), "Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds", *Financial Analysts Journal*, Vol. 49, No. 6, pp. 62-66.

Henke, H. (2016), "The effect of social screening on bond mutual fund performance", *Journal of Banking & Finance*, Vol. 67, pp. 69-84.

Hendricks, D., Patel, J. and Zeckhauser, R. (1993), "Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988"*, The Journal of finance (New York),* Vol. 48, No. 1, pp. 93-130.

Ibikunle, G. and Steffen, T. (2017), "European Green Mutual Fund Performance: A Comparative Analysis with their Conventional and Black Peers", *Journal of Business Ethics*, Vol. 145, No. 2, pp. 337-355.

Jensen, M.C. (1968), "The Performance of Mutual Funds in the Period 1945–1964", *The Journal of Finance*, Vol. 23, No. 2, pp. 389-416.

Kaur, I. (2018), "Effect of mutual funds characteristics on their performance and trading strategy: A dynamic panel approach", *Cogent Economics & Finance*, Vol. 6, No. 1, pp. 1493019.

Kjetsaa, R. and Kieff, M. (2016), "Impact of Expenses, Turnover and Manager Tenure on Blend Fund Performance"*, Journal of Business and Accounting,* Vol. 9, No. 1, pp. 99-115.

Kleiman, R.T. and Jun, K.W. (1988), "The Size of Mutual Funds and Risk-Adjusted Performance: Some New Evidence"*, Journal of Applied Business Research,* Vol. 4, No. 3, pp. 50-55.

Klir, G.J. and Yuan, B. (1995), *Fuzzy sets and fuzzy logic : theory and applications.* Prentice Hall, Englewood Cliffs (NJ).

Kumbure, M.M., Tarkiainen, A., Luukka, P., Stoklasa, J. and Jantunen, A. (2020), "Relation between managerial cognition and industrial performance: An assessment with strategic cognitive maps using fuzzy-set qualitative comparative analysis", *Journal of Business Research*, Vol. 114, pp. 160-172.

Lemak, D., and P. Satish. "Mutual Fund Performance and Managers' Terms of Service: Are There Performance Differences?" *The Journal of Investing*, Winter 1996, pp. 59-63.

Lintner, J. (1965), "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets"*, The Review of Economics and Statistics,* Vol. 47, No. 1, pp. 13-37.

Markowitz, H. (1952), "Portfolio Selection", *The Journal of Finance* (New York), Vol. 7, No. 1, pp. 77.

Morningstar (2015), *How to Use the Sharpe Ratio.* [Web document] [Accessed 17.4.2021]. Available: http://news.morningstar.com/classroom2/course.asp?docId=2932&page=5&CN=com

Morningstar (2016), *Morningstar Introduces Industry's First Sustainability Rating for 20,000 Funds Globally, Giving Investors New Way to Evaluate Investments Based on Environmental, Social, and Governance (ESG) Factors*, "Mar 1, "-last update. [Web document] [Accessed 21.2.2021]. Available: https://newsroom.morningstar.com/newsroom/news-archive/press-release-details/2016/Morningstar-Introduces-Industrys-First-Sustainability-Rating-for-20000-Funds-Globally-Giving-Investors-New-Way-to-Evaluate-Investments-Based-on-Environmental-Social-and-Governance-ESG-Factors/default.aspx

Morningstar Nordic Info Team (2020), "Apr 30, "-last update, Morningstar ostaa koko Sustainalyticsin. [Web document] [Accessed 21.3.2021]. Available: https://www.morningstar.fi/fi/news/201999/morningstar-ostaa-koko-sustainalyticsin.aspx

Morningstar Office (2021), Manager Tenure. Glossary Definitions. [Web document] [Accessed 5.4.2021]. Available: http://awgmain.morningstar.com/webhelp/glossary_definitions/mutual_fund/mfglossary_Manager_Tenure.html

Morningstar Research (2019), "Morningstar Sustainability Rating Methodology", Morningstar. [Web document] [Accessed 20.3.2021]. Available: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/744156_Morningstar_Sustainability_Rating_for_Funds_Methodology.pdf

Nagy, Z., Kassam, A. and Lee, L. (2016), "Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies", *The Journal of Investing,* Vol. 25, No. 2, pp. 113-124.

Nofsinger, J. and Varma, A. (2014), "Socially responsible funds and market crises", *Journal of Banking & Finance*, Vol. 48, pp. 180-193.

OECD (2020), OECD Business and Finance Outlook 2020: Sustainable and Resilient Finance, OECD Publishing, Paris, [Web document] [Accessed 5.4.2021]. Available: https://doi.org/10.1787/eb61fd29-en

Perold, A. F. and Salomon. R.S. (1991), "The Right Amount of Assets under Management", *Financial Analysts Journal,* Vol. 47, No. 3, pp. 31-39.

Ragin, C. C. (2008), Redesigning Social Inquiry: Fuzzy Sets and Beyond. University of Chicago Press.

Riedl, A. and Smeets, P. (2017), "Why Do Investors Hold Socially Responsible Mutual Funds?", *The Journal of Finance (New York),* Vol. 72, No. 6, pp. 2505-2550.

Sharpe, W.F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk"*, The Journal of Finance (New York),* Vol. 19, No. 3, pp. 425-442.

Sharpe, W.F. (1966), "Mutual Fund Performance", *The Journal of Business* (Chicago, Ill.), Vol. 39, No. 1, pp. 119-138.

Smit, E. v. d. M. and Friis, L.B. (2004), "Are some fund managers better than others? Manager characteristics and fund performance", *South African Journal of Business Management*, Vol. 35, No. 3, pp. 31-40.

Sossong, A. (2014), Institutional asset management: how are professional investors affected by legislature and corporate governance? Translated by Anonymous. Anchor Academic Publishing, Hamburg, Germany.

Statista Research Department (2020), Number of mutual funds worldwide 2009-2019. [Web document] [Accessed 5.4.2021] Available: https://www.statista.com/statistics/278303/number-of-mutual-funds-worldwide/

Statista Research Department (2021), Mutual funds - statistics & facts. [Web document] [Accessed 4.4.2021] Available: https://www.statista.com/topics/1441/mutual-funds/#dossierSummary__chapter2

Steen, M., Moussawi, J.T. and Gjolberg, O. (2020), "Is there a relationship between Morningstar's ESG ratings and mutual fund performance?", *Journal of Sustainable Finance and Investment*, Vol. 10, No. 4, pp. 349-370.

Stoklasa, J., Luukka, P. and Talášek, T. (2017), "Set-theoretic methodology using fuzzy sets in rule extraction and validation - consistency and coverage revisited", *Information Sciences*, Vol. 412-413, pp. 154-173.

Stoklasa, J., Talášek, T. and Luukka, P. (2018), "On consistency and coverage measures in the fuzzified set-theoretic approach for social sciences: Dealing with ambivalent evidence in the data." In Proc. of the 36th Inter. Conf. on Mathematical Methods in Economics, pp. 521–526.

Treynor, J.L. (1965), "How to Rate Management of Investment Funds", *Harvard Business Review*, Vol. 43, No. 1, pp. 63–75.

Tufano, P. and Sevick, M. (1997), "Board structure and fee-setting in the U.S. mutual fund industry", *Journal of Financial Economics,* Vol. 46, No. 3, pp. 321-355.

US SIF. (2020), “2020 Report on US Sustainable and Impact Investing Trends.” US SIF Foundation. [Web document] [Accessed 21.3.2021] Available: https://www.ussif.org/files/Trends%20Report%202020%20Executive%20Summary.pdf

US SIF. (n.d.) Sustainable Investing Basics. [web document] [Accessed 18.4.2021] Available: https://www.ussif.org/sribasics

Wang, C.D., Chen, Z., Lian, Y. and Chen, M. (2020), "Asset selection based on high frequency Sharpe ratio", *Journal of Econometrics*.

Yan, X. (2008), "Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance", *Journal of Financial and Quantitative Analysis,* Vol. 43, No. 3, pp. 741-767.

Yin-Ching, J. and Mao-Wei, H. (2003), "Mutual fund attributes and performance", *Financial Services Review*, Vol. 12, No. 2, pp. 165-178.

Zadeh, L.A. (1965), "Fuzzy sets", *Information and Control,* Vol. 8, No. 3, pp. 338-353.

Zimmermann, H.- (2010), "Fuzzy set theory", *Wiley interdisciplinary reviews. Computational statistics,* Vol. 2, No. 3, pp. 317-332.