ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Msc Financial Economics

Doing good by doing well: A closer look at the sustainable US open-end mutual fund market

Name student: Nils Eijsvogel

Student ID number: 455275

Supervisor: E. Smajlbegovic

Second assessor: XXX

Date final version: November 2022

Doing good by doing well: A closer look at the sustainable US open-end mutual fund market

Nils Eijsvogel

November 2022

Abstract

This study examines the performance of US open-end mutual funds, relative to their Morningstar Sustainability Rating, over the period January 2000 – February 2021. First, this paper applies the (1) Capital Asset Pricing Model (CAPM), (2) the Fama and French (1993) three-factor model, and (3) the Carhart model to evaluate the potential outperformance of the funds relative to their passive benchmark. Second, this paper observes multiple performance measurement statistics in order to rank the different sets of mutual funds on their performance. Finally, this paper looks at the dependence of sensitivity of mutual funds flows on a set of explanatory variables and fund characteristics. The main findings do not indicate a significant outperformance of sustainable mutual funds relative to their non-sustainable peers. However, they do not significantly underperform as well. Thus, investors could invest in sustainable investments without harming their performance.

Contents

1 Introduction

Sustainable innovations become more important globally, since the environmental concerns keep rising. As a united front, each part of the society should start to think more about the environment. Naturally, this also includes companies. Institutions start to make their operations more sustainable or innovate more sustainable products. This increasing sustainable trend also reflects on financial markets. Moreover, sustainable investing becomes more important in the financial markets as well. Since companies focus more on sustainable decision-making, investors should pay more attention to sustainable oriented companies. This upward trend also contribute positively to your investments, since sustainable investing could increase the growth opportunities for your company (Rigby and Tager (2008)), and, therefore, improve the investor's portfolio.

One problem that occurs when talking about sustainable investing, and sustainability in general, is that there is no general definition of sustainability. For instance, one defines it as the capability to remain at a particular level or point in time (Gruen et al. (2008)), where the other comes up with a more extensive definition. Moreover, to cite J. E. Moore et al. (2017), they define sustainability as: *"(1) after a defined period of time, (2) the program, clinical intervention, and/or implementation strategies continue to be delivered and/or (3) individual behavior change (i.e., clinician, patient) is maintained; (4) the program and individual behavior change may evolve or adapt while (5) continuing to produce benefits for individuals/systems"*. Missing a general definition of sustainability complexes the concept, causing a negative effect on the research regarding sustainability.

The term Corporate Social Responsibility (CSR) is a well-known principle in the upward sustainability trend among corporations. Sanie and Durres (2015) describes this term as the improvements of ethics inside institutions. In particular, this means an increase in responsibility to take ethical actions and to keep making these ethical actions. This introduces the start of a new era, in contrary of the neoclassical economists (Friedman (1970)). Where the neoclassical economist only focuses on maximizing its profits, CSR focuses more on the responsiveness and creativity of its corporation to battle global warming concerns. Luckily, the incorporation of CSR in businesses also has a positive effect, where it increases the performance of the business (Beckman, Colwell, and Cunningham (2009); Heyward (2020); Palmer (2012)).

The inclusion of sustainability in the corporate sector is an important factor in financial markets as well. Moreover, in the report of Alliance (2018), the growth in Social Responsible Investing (SRI) finance becomes visible. This report states that more than $30 trillion worth of investments are made in responsible investments. The basic principle of Social Responsible Investing is to avoid investing in sin stocks as a investor, also called negative screening. Stocks are referred to as "sin" when the company is operating in "sin" industries, such as air travel, gambling or tobacco. Whereas SRI becomes larger, the corresponding financial performance is research more intensively. However, the conclusions of these papers are not heading into the same direction. For instance, one example is the "doing good but not well" hypothesis. This hypothesis state that an increase in sustainable investments could decrease the financial performance of those investments (Statman and Glushkov (2009)), because of the higher expected returns of non-sustainable stocks. Fortunately, most research find a positive relationship between Social Responsible Investing and financial performance (Henisz, Koller, and Nuttall (2019); Unruh et al. (2016); Gompers, Ishii, and Metrick (2003)). Other papers look more closely at the volatility of those investments, where they find evidence for a reduction of the volatility when investing in sustainable stocks (Morelli and D'Ecclesia (2021); S. Bodhanwala and R. Bodhanwala (2019)).

A part of the financial market where sustainable investing is growing, is in the mutual fund market. A mutual fund is a basket of financial securities actively supervised by the fund manager. The first ever mutual fund is introduced in the Netherlands, where a Dutch broker invented a trust, called *"Eendragt Maakt Magt"*. This way, investors could obtain a diversified portfolio for a relatively small investment. As an individual investor, you can collect a small part of this fund, to price efficiently obtain a well diversified investment in your portfolio. Although the individual investors pay a management fee, in general, these fees do not outweigh the benefits of an active management (Avramov and Wermers (2006)). Nowadays, mutual funds become so popular that, in the beginning of the 21^{st} century, there are more mutual funds listed on the New York Stock Exchange than stocks (Factbook (2002)). Moreover, just as the shift to sustainability in the rest of the world, the same upward trend is going on in the mutual fund market. Research of UNCTAD (2021) reveals that the number of sustainable mutual funds doubled in the last five years. Another research finds that the amount of new outflows for sustainable mutual funds is twice as big as the outflow for their less sustainable peers (Hartzmark and Sussman (2019)). One reason behind this upwards trend is that investors more often combine their values with their investments (Prior (2015)). Another reason is that there is an increase in female fund managers (Bliss and Potter (2002)), and that these female fund managers more often choose for a sustainable investment than their male peers (Management (2021)). Although there is literature that support the convex relationship between mutual fund flows and past performance (Sirri and Tufano (1998); Huang, Wei, and Yan (2007); Ferreira et al. (2012)), sustainable investors care less about performance when investing in sustainable investments (Bollen (2007)). Therefore, the inflows of sustainable-oriented funds could be less sensitive to past performance then their non-sustainable peers.

This paper combines previous literature on sustainable investing and the performance of mutual funds. The literature about the possible increase in performance when there is a focus shift to sustainable investing, together with the increase of interest in sustainability of mutual fund managers, could potentially lead to an outperformance of mutual funds that focus on sustainable investing relative to mutual fund that do not focus on the sustainability of their investments. However, since the previous literature partly disagrees with this outperformance, it still not clear whether an increase in the sustainability of mutual funds is equivalent to an increase in financial performance. This paper investigates and compares the performance of sustainable mutual funds and non-sustainable mutual funds. Therefore, the research question is stated as follows:

"Do US open-end mutual funds that construct a sustainable portfolio outperform their less sustainable peers, between January 2000 and February 2021?"

This paper investigates the research question by via multiple directions. First, the potential outperformance of mutual funds relative to the fund's benchmark could improve when the fund is focusing more on sustainable investments. Second, this paper looks at several performance statistics to evaluate the performance of the two different sets of mutual fund. Finally, more sustainable mutual funds could attract larger relative flows than their non-sustainable peers. The paper refers to sustainable mutual funds and their less sustainable peers as high rated funds and low rated funds, respectively. Therefore, the corresponding hypotheses are:

- H_1: High rated US open-end mutual funds more often outperform their relative benchmark than low rated US open-end mutual funds between January 2000 and February 2021.
- H_2: High rated US open-end mutual funds outperform low rated US open-end mutual funds according to various performance statistics between January 2000 - February 2021.
- H_3: High rated US open-end mutual funds flows are less sensitive to past performance than low rated US open-end mutual funds flows between March 2018 - February 2021.

The mutual fund data for this research come from the Morningstar database. This dataset regards monthly returns for US open-mutual fund data over the period January 2000 - February 2021. The mutual funds are split into two sets. The first set of mutual funds scores the highest ($\star\star\star\star\star$) in terms of the Morningstar Sustainalytics rating. The other set of mutual funds scores the lowest ($\star$) in terms of the same rating. Therefore, this paper refers to these mutual fund sets as high rated funds and low rated funds. Moreover, the risk factor data come from the Kenneth R. French library website. This data regard monthly returns for the market factor (MKT-R_f), the size factor (SMB), the value factor (HML), the momentum factor (MOM), and the risk-free rate (R_f) over the same sample period. The data for the mutual funds flows also come from the Morningstar database. Finally, the Total Net Assets (TNA) data also come from the Morningstar database. This data set regards monthly Total Net Assets (TNA) for US open-end mutual funds ranging from March 2018 - February 2021.

First, this paper investigates the performance of the mutual funds by constructing performance models. The performance models are the same models as in the paper of Eling and Faust (2010), and include the Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model, and the Carhart (1997) model. Then, the alphas and the adjusted R^2 of the models using the different sets of mutual funds serve as the measurements of the performance. In particular, the alphas of the model present the potential outperformance of the mutual fund relative to a passive benchmark. Looking at the R^2, there is no significant outperformance of the high rated funds versus the low rated funds. Additionally, the alpha distribution of the performance models does not find a significant indication that high rated fund managers more often outperform their relative benchmark then the low rated fund managers. A robustness check that split the overall sample period into five subperiods obtains evidence that the performance models are not robust for different sets of data, since the alpha distribution is different for the five subperiods. All together, there is no significant indication that high rated funds more often outperform their passive benchmark relative to low rated funds.

Second, this paper looks at various performance measurement statistics to further evaluate the performance of the US open-end mutual funds, It also looks at the performance measurement statistics of the risk factors, that serve as passive benchmark indices for the different fund sets. This paper finds no significant preference for the set of more sustainable mutual funds over their less sustainable peers.

Finally, this paper observes the mutual fund flows of the US open-end mutual funds. It does so by regressing the relative flows on fund characteristics and other explanatory variables, including a performance measurement and a dummy variable. Moreover, this paper chooses Jensen (1968) alpha as the measurement for performance and the size and age of the funds as characteristics. The regression shows no significant coefficients for the explanatory variables. This means that there is no significant evidence that high rated funds are less sensitive to performance than low rated funds.

The main findings indicate no significant outperformance of either of the different sets of mutual funds. Therefore, this paper rejects all its hypotheses. Investors do not increase their overall performance when they increase their sustainable investments. However, to look at the positive side, investors also do not decrease their overall performance when they increase their sustainable investments. This could indicate that investors should focus more on the sustainability factor in their investment, as it is better in terms of environment, and it does not harm their financial performance.

2 Literature Review

2.1 Sustainability

The concerns about climate change are a main topic in today's society. However, problems concerning global warming started much earlier than some may think. In particular, Abram et al. (2016) state that emission gas started to heat up the world already around the 1830s. However, the significant increase of the concerns about global warming started around 2012, since research show that there is a critical point concerning global warming at around this year (Metz, Hulme, et al. (2013); Olmstead and Stavins (2012)). Consequently, at around the same time, governments introduce regulations against these global warming issues (such as the post-2012 climate policy of the EU). These particular regulations could reduce the global warming, since research show that most of the environmental issues nowadays are caused by human activities (Dutta (2017);

Steg and Vlek (2009); Hertwich and Peters (2009); Swim, Clayton, and Howard (2011)).

Due to the global warming concerns, the need for the inclusion of sustainable decision-making in our day-to-day life is huge. To really understand the principal of sustainability, the concept needs to be properly explained first. However, research show that there is no unique definitions of sustainability and that there is no common use of the principal in the existing literature (Proctor et al. (2015)). One example of a, relatively simple, definition of sustainability is the one of Gruen et al. (2008), where they state that sustainability is equal to the capability to remain at a certain level or a certain point. A more extensive definition comes from J. E. Moore et al. (2017), where they divide their definition into five different components: *"(1) after a defined period of time, (2) the program, clinical intervention, and/or implementation strategies continue to be delivered and/or (3) individual behavior change (i.e., clinician, patient) is maintained; (4) the program and individual behavior change may evolve or adapt while (5) continuing to produce benefits for individuals/systems"*. Regarding global warming, sustainability is a concept where, due to human behavioral change, the state of the climate should be maintained or even improved during time.

Since the introduction of sustainable decision-making is becoming more important in every industry, businesses that incorporate sustainability in their company are attracting the attention nowadays as well. This means that investors seek shares of companies that include sustainability in their business. Therefore, the importance of incorporating sustainability increases. However, the question is if this new sustainable trend is also profitable for companies as well as for individual investors.

2.1.1 Sustainability in the corporate sector

As stated in Section 2.1, human activity is the main factor for environmental issues. Therefore, the largest sustainable injections should be made in areas with this kind of human activity, such as the corporate sector. Nowadays, governments, as well as other parts of the society, expect companies to aim their business in line with improvements of the environment. Therefore, Corporate Social Responsibility (CSR) becomes more important in the society. The basic idea of CSR is that companies improve their ethics, in such a way that they hold their businesses responsible to take ethical actions and obligate their selves to stick to these measures (Sanie and Durres (2015)). Just as in other areas of the world, the critical point for Corporate Social Responsibility lies around the year of 2012. For example, Visser (2011) states that, before 2011, the inclusion of social responsibility fails to have significant impact. This includes a new era after that year. Although the incorporation of CSR contradicts the neoclassical economic paradigm, because it can contradict with obtaining maximizing profits (Friedman (1970)), corporations should now also focus on creativity, scalability, responsiveness and circularity, to really battle the global warming issues. Moreover, Schiller (2015) even state that companies should incorporate CSR, even though it could potentially cost money over the long-run. In his theory, he appoints a third phase of the relationship between companies and the society, where he points out the importance of the so-called "shared value", instead of the individual value of each company.

Since the rise of the Corporate Social Responsibility, companies obtain a score on how they perform in terms of sustainability. A well-known score is the Environmental, Social and Governance (ESG) score. ESG scores assign three separate scores for the Environmental (E), Social (S) and Governance (G) aspect of the company, as well as an overall rating of the three areas. This way, investors can now take this new firm characteristic into account in their decision-making process. Sadly, there are some inconsistencies in the ESG rating process, since there is only a small group of firms that assign these scores to corporations, which could lead to some limitations (Kotsantonis and Serafeim (2019)).

Fortunately, research show that prioritizing CSR often goes hand in hand with an increase in performance. For instance, Beckman, Colwell, and Cunningham (2009) show that activities concerning CSR can create shareholder value. Their remark on this is that managers need to be monitored properly by long-term investors. Next, Heyward (2020) finds more possible benefits of CSR, namely the increase of accountability to investors and the possibility to save more money. Other positive side effects are that it can improve the relationship and the engagement with existing customers, and it can draw new customers to your company. Additionally, research show that customers are prepared to pay a premium for products and or services of companies that incorporate CSR in their business relative to their non-CSR peers. This may be the cause of an increase in gross margin (Palmer (2012)).

Furthermore, Tsoutsoura (2004) research if there is a relationship between Corporate Social Responsibility and the financial performance of the company, and whether the potential effects are positive or negative. In their paper, they find a significant positive effect of CSR on financial performance, and that this effect is mainly the cause of bottom-line benefits. Next, there is also research that investigate the possible relationship with CSR and certain risk factors. Moreover, CSR could help diminish global supply chain risk (Cruz (2013)). Finally, Diemont, K. Moore, and Soppe (2016) and Lööf, Stephan, et al. (2019) show that there is a significant reduction of downside tail risk when a company is focused on their Corporate Social Responsibility.

Nevertheless, there is a disadvantage in the rise of assigning ESG scores to firms. Moreover, firms that perform poorly could use ESG as a cover for their underperformance. In the paper of Flugum and Souther (2021) they restate a finding in the paper of Ryan Flugum of the University of Northerm Iowa and Matthew Souther of the University of South Carolina. They find that the underperformance of a company positively correlated with the amount that the respective manager publicly embraces their ESG priority. This could eventually lead to a chain reaction, where managers of sustainable funds invest in so-called ESG-focused firms, and therefore indirectly invest in companies with poor performances.

2.1.2 Sustainability in the financial markets

The inclusion of sustainability in the corporate sector and the corresponding ESG score of a company are important factors in financial markets as well. In particular, the investor's perspective of the effect of ESG and sustainability in their investments is important. Partly because of that fact, the fraction of responsible investments is rising during the last years. Moreover, in 2018, more than $30 trillion worth of investment lies in this area of the market (Alliance (2018)). The same report shows that the growth in sustainable finance rises more rapidly in the US than in Europe. Another reason is that millennials are one of the driven forces of the rise in sustainable investments. Moreover, Adamczyk (2021) state that, because millennials are more searching for growth opportunities than other investors, they find firms with high ESG scores more interesting. On the contrary, Akerlof (1980) and Becker (2010) find that individual put their possible financial gain over sustainable investing.

Whereas firms try to avoid sin activities, such as avoiding emitting huge amounts of emissions, in favor of their ESG scores, investors try to avoid sin stocks, also called negative screening. This is the basic principle of Social Responsible Investing (SRI) (Liang and Renneboog (2020)). In their paper, they name two other generations of SRI investors: the ones that include companies that prioritize their sustainable operations (positive screening), and the ones that combine both positive and negative screening. However, the latter group of investors seems the most common generation, since investors favor a more holistic investment strategy rather than an exclusionary approach. Investors have a preference for including firms with high ESG scores than excluding their less ESG-related peers (Berry and Junkus (2013)). Furthermore, Riedl and Smeets (2017) find that both social preferences and social signaling explain Socially Responsible Investment decisions. Financial motives play less of a role.

However, short sellers, as informed investors, do not participate in the upward trend of sustainable investing. Moreover, A. Jain, P. K. Jain, and Rezaee (2016) investigate whether short sellers take the Corporate Social Responsibility and ESG performance of a company into account. Unfortunately, they find that short sellers are more interested in firms with a low ESG score than firms with a high ESG score. Therefore, there is a negative correlation between the ESG score of a firm and the probability that short seller target that company. Rusinova and Wernicke (2019) finds the same result. That paper finds strong evidence for its hypothesis that managers increase the firm's CSR performance when they face a growth in short-sellers interest.

2.1.3 Sustainability and financial performance

Although the focus of investors shifts more towards sustainable investing, they still need to focus on the profitability of their investments as well. Moreover, there is a "doing good but not well" hypothesis. This hypothesis state that investors who embrace social responsibility in their investment decisions find a negative effect on their financial performance. Statman and Glushkov (2009) find support for this hypothesis. They argue that sin stocks - such as tobacco, airlines, gambling, and pollution - obtain a higher expected return. So, eliminating the sin stocks in your portfolio for the sake of the environment could damage your expected return, and thus the performance of your investment.

However, in theory, investing in ESG could actually increase the value of your investment (Henisz, Koller, and Nuttall (2019)). Moreover, Unruh et al. (2016) state that corporate sustainability significantly correlates with financial performance. This is in line with the research of Gompers, Ishii, and Metrick (2003). The research of Paul Gompers of the Harvard Business School in 2003 introduces the topic of corporate performance and ESG investing as one of the first, which makes it one of the prominent papers in this area. In this paper, they find a significant positive relationship between corporate sustainability and their relative performance.

Since then, a lot of research focuses on the relationship. For example, the paper of Clark, Feiner, and Viehs (2015) investigate over 200 different historical research sources, and state that "80% of the reviewed studies demonstrate that prudent sustainability practices have a positive influence on investment performance". This is in line with the research of Chava (2014). In addition, the Asset and Wealth Management division of the Deutsche Bank, together with the University of Hamburg, conduct a survey to further research this issue. They even look at the entire universe of academic papers of 2,250 studies between 1970 and 2014. They come up with a slightly lower percentage of 62.6% in favor of the outperformance of ESG related investments, and they find a percentage of less than 10% where ESG has a negative impact on the performance. Derwall et al. (2005) find a similar result. They state that firms that support eco-friendly business perform better than their peers.

There is also research available that focuses on the volatility differences of ESG-related investments relative to non-ESG investments. Moreover, Morelli and D'Ecclesia (2021) zoom in on the Environmental (E) part of ESG investing and compare European portfolios that contain a high E score with their counterparts. In this paper, the volatility of portfolios with high E score stocks is significantly lower than their peers, mainly in times of financial distress. However, there is literature available in the research of the relationship of incorporating sustainability and the corresponding financial performance that do not find a positive relationship. In particular, Xiao et al. (2013) find no support in favor of sustainable investing in terms of financial performance.

Furthermore, there is even literature available that find a negative effect on financial performance when focusing on ESG-related investments. For instance, Chava (2010) finds evidence that stocks that are labeled as "sin", and thus excluded from the socially responsible investors, have a significant higher expected return than their more sustainable counterparts. This is in line with the research of Riedl and Smeets (2017), where they state that Socially Responsible Investing earns lower relative returns than conventional investing. 'Relative to SRI funds, they also state that SRI funds demand higher management fees, which both could negatively influence the decision-making of social responsible investors. Adler and Kritzman (2008) find that the costs for socially responsible investing exceeds the costs for their counterparts as well. Finally, although S. Bodhanwala and R. Bodhanwala (2019) agrees with Morelli and D'Ecclesia (2021) about the significant lower volatility for sustainable investments, they argue that this goes hand in hand with an underperformance relative to their benchmark.

The conclusion about this topic are divided. Hong and Kacperczyk (2009) state that this is due to the divergent definition of the concept of sustainable investing. As well as the paper of Liang and Renneboog (2020), the paper of Hong and Kacperczyk (2009) make a distinction between investors that include companies in line with their beliefs, investors that exclude companies that are "sin", and investors that conduct both principles.

2.2 Mutual funds

There are many ways to participate in financial markets. For example, investors could construct a portfolio of stocks and bonds. However, especially for households investors, this may not be the most efficient way to participate in financial markets. Therefore, investing in mutual funds could be a more efficient solution for households. Moreover, in 2019, 45.5% of the individuals and households in the United States invest in mutual funds (SIFMA (2021)).

A mutual funds is a portfolio, where a manager actively remains supervision op the portfolio. Moreover, as an individual investor, you do not have to own the entire mutual fund. Individual investors collect a piece of the pie, which eventually makes up for the total value of the mutual fund. Therefore, for a small investment, you can already obtain the benefits of a diversified portfolio. Although investors need to compensate the manager costs, active management of the fund still adds significant value (Avramov and Wermers (2006)).

There are three different types of mutual funds (Elton and Gruber (2013)). The first one are open-end funds. Managers of open-end funds are not limited in the number of shares they can issue. In contrary, closed-end funds are limited in the number of shares it can issue. Managers of closed-end funds can issue new shares only via an Initial Public Offering (IPO). Finally, Exchange-traded funds (ETF) are the last type of mutual funds. Exchange-traded funds are funds that trade on major stock exchanges. The difference between ETFs and the other mutual funds is that the price of an ETF fluctuates the entire day, whereas the price of the other funds is adjusted at the end of every day, depending on its Net Asset Value (NAV). Another difference is that the manager of an ETF construct the funds in such a way that it tracks a major index, such as the S&P 500 index.

The origin of mutual funds lies in the Netherlands. Rouwenhorst (2004) state that during the second half of the 18^{th} century, a Dutch broker introduced a construction where he invites investors to fund a piece in a trust named "*Eendragt Maakt Magt*", to enjoy a diversified portfolio for a small-scale investment. The amount of mutual funds continue to grow since. In particular, during the 1990s, the amount of mutual funds extend in most countries over the world. Klapper, Sulla, and Vittas (2004) names two main reasons for this expansion: the change in attitude of investors in the financial market and the corresponding rise of confidence in the financial market. In fact, in the beginning of the 21^{st} century, the number of mutual funds surpasses the amount of stocks that are listed on the New York Stock Exchange (Factbook (2002)).

2.2.1 Sustainable mutual funds

Consistent with the shift in the financial market to sustainable investing, mutual fund managers construct their portfolio with more sustainable securities as well. Moreover, research of UNCTAD (2021) finds an increase in sustainable mutual funds, especially in the last few years. In particular, they find that the number of these funds doubled in the last five years. Therefore, the number of sustainable mutual fund in the market almost reached the 4,000 in June 2020. They also find numbers for the net investment flows of mutual funds. Moreover, this total number reached $159 billion in 2019, with the expectation that this number will exceed the $300 billion in 2020. The research of Hartzmark and Sussman (2019) find the same increase in sustainable mutual funds. They state that investors in the market nowadays value sustainability. While the value of net outflows for low sustainability mutual funds is around $12 billion, the value of net outflows for high sustainability mutual funds is double this value.

Similar to Section 2.1, although the numbers increase more rapidly in the last few years, the urge for more sustainable investment in the mutual fund market was firstly mentioned a long time ago. In particular, Moskowitz (1972) was one op the first researchers that fires up the need for an increase in sustainable investments in the mutual fund market. Not only in terms of the positive social effect, but also in terms of the possible positive impact on the performance of the funds.

The reasons for this increase of the number of sustainable mutual funds in the market are diverse. For instance, investors more often "put their money where their mouth is". In terms of investments, they more often "put their money where their values are" (Prior (2015)). This means that if investors have concerns about the climate, they include this in their investment decisions as well. Since the concerns about climate change is a global upward trend, this trend also shows up in the financial market. Another upward trend that is also increasing in the financial market is the increase of female investors. The amount of females in the financial market is rising, both as individual investors and institutional investor. In particular, Bliss and Potter (2002) state that the percentage of female fund managers is twice as high as the percentage five years ago. They also state that this percentage is going to increase even more in the upcoming years. Since women, in general, have different character traits as men, this could potentially influence the investor decision-making in the market. A recent study by Management (2021) draws the same conclusion. Moreover, the research shows that female fund manager are twice as likely to find Social Responsible Investing important in their investment decision-making as male fund managers.

Another difference with conventional fund managers, is the willingness to take risks (Dantas (2021)). The paper suggests that the preferences of ESG investors affect ESG managers' willingness to engage in opportunistic trading without significantly affecting their real portfolio decisions. Finally, Dantas (2021) also finds that ESG fund managers have longer investment horizons than their peers.

2.2.2 Mutual funds and financial performance

After making a distinction between mutual funds in general and sustainable mutual funds, this paper look at the financial performance of mutual funds. Research finds a positive persistence of mutual fund performance (Hedricks, Patel, and Zeckhauser (1993); William and Ibbotoson (1994); Brown and Goetzmann (1995); Wermers (1997)). This evidence is from a short term horizon and is mostly due to the "hot hand" effect, a cognitive bias where individuals strongly believe in the momentum effect of their securities. Moreover, Grinblatt and Titman (1992) finds that the performance persistence of mutual funds are positively correlated with the abnormal returns obtained by mutual funds. However, Carhart (1997) only finds this performance persistence by the worst performing mutual funds. Moreover, recent studies show that, on average, mutual funds do not outperform their relative benchmarks. In particular, Risk (2022) states that funds do not outperform their benchmarks, especially when the funds are stamped active. Additionally, Chen et al. (2004) finds evidence for a negative relationship between fund size and the financial performance.

An important fund characteristic that impact the investment decision of individual investors in the rate of sustainable investing of the mutual fund. Moreover, retail investors shift their investment from mutual funds that do not focus on sustainable investing to mutual funds that do focus on sustainable investing (Ammann, Bauer, Fischer, Mueller, et al. (2017); Ammann, Bauer, Fischer, and Müller (2019)). One way to measure the sustainable investment rate of a mutual fund, is to look at the Morningstar Sustainability Rating. The Morningstar Sustainability Rating is a measurement that points out the degree of sustainable investing of a portfolio or fund. Moreover, it looks at the holdings of the portfolio of fund and measure their relative ESG risks (Justice and Hale (2016)). Research finds investors argue that Morningstar is the best and most prominent ranking service (Sharpe (1998); Blume (1998); Blake and Morey (2000)). Furthermore, a high Morningstar Sustainability Rating could influence the inflows of the particular mutual fund. Moreover, whereas the alphas of high sustainable funds do not significantly differ from the alphas of low sustainable funds (Dolvin, Fulkerson, and Krukover (2019)), there is strong evidence that the Morningstar Sustainability Rating contribute positively to the inflows of the mutual fund (Hartzmark and Sussman (2019); Del Guercio and Tkac (2008)).

2.2.3 Mutual fund flows

As mentioned earlier, the most important benefit for an individual investor to invest in a mutual fund, is to obtain a piece of a diversified portfolio for a relatively small investment. However, this is only a benefit for the individual investor when the mutual fund is performing well. Therefore, the performance of mutual funds is important for the liquidity inflows of a mutual market. Fortunately, Warther (1995) states that there is a positive correlation between fund inflows and market returns. Goriaev, Nijman, Werker, et al. (2002) finds the same relationship, where, in his research, better performing mutual funds attract higher relative flows.

Moreover, Goriaev, Nijman, Werker, et al. (2002) states that the relative flows depend on more than the performance of the mutual fund. Additionally, smaller and younger funds also attract higher relative flows (Chevalier and Ellison (1997)).

The convex relationship between mutual fund flows and past performance that Goriaev, Nijman, Werker, et al. (2002) find, is a familiar relationship in the mutual fund literature (Sirri and Tufano (1998); Huang, Wei, and Yan (2007); Ferreira et al. (2012)). However, some literature find a linear relationship between the fund flows and past performance (Spiegel and Zhang (2013)). They state that convexity occur when the model is misspecified. Many researchers support the convex relationship, since conventional investors chase past performance, which increases fund flows (Rohleder (2015)).

Where conventional investors chase past performance, sustainable investors care less about performance when investing in sustainable investments (Bollen (2007)). Therefore, the inflows of sustainable-oriented funds could be less sensitive to past performance then their non-sustainable peers. Although sustainable mutual fund inflows should be less sensitive to past performance, literature find shift in inflows from non-sustainable funds to sustainable funds. Therefore, there should be a positive relationship between inflows and performance. Since inflows increase, performance should increase as well. (Hartzmark and Sussman (2019); Del Guercio and Tkac (2008)).

3 Data

This paper uses different data sets containing data for US open-end mutual funds and Fama and French (1993) factors. The data sets for the main research range between January 2000 and February 2021.

First, this paper obtains data from the Morningstar database. In this database, this paper obtains monthly returns for US open-end mutual funds over the given time range. To compare the data sets of sustainable mutual funds and its non-sustainable peers, there needs to be a distinction whether a mutual fund is related to sustainability or not. This data is received from the Morningstar database, where they stamp sustainable ratings (Sustainalytics) to funds, similar to ESG ratings. Therefore, for the set of sustainable mutual funds, this paper picks the mutual funds that scores highest rate ($\star\star\star\star\star$) in terms of the Morningstar Sustainalytics rating. Next, for the set of their non-sustainable peers, this paper picks this set of mutual funds as the mutual funds that contains a rate of ($\star$) in terms of the Morningstar Sustainalytics rating. This boils down to a data set of a total of 724 US open-end mutual funds, split into 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating, and 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating. From now on, this paper refers to these funds as high (rated) funds and low (rated) funds, respectively.

Second, this paper obtains data from the Kenneth R. French library website. At this website, this paper acquires monthly data of the Fama and French (1993) 3-factor model over the research period. The first factor

$(MKT - R_f)$ represents the excess market return. The other two factors, (SMB) and (HML), represents the size and value factor, calculated by the firm's market capitalization and book-to-market ratio, respectively. Finally, this paper monthly data of the MOM factor from the same website, capturing the momentum effect. Table 1 in the Appendix shows the descriptive statistics of the mutual fund returns and the returns for the four factors.

3.1 Fund flow data

Despite that the overall research period ranges from January 2000 - February 2021, the Total Net Assets (TNA) for the same US open-end mutual funds are only available from March 2018 - February 2021 on monthly basis. Therefore, in this part of the research, the research period is set from March 2018 - February 2021. Moreover, during this period, the Morningstar database does not collect all the monthly TNA for all the mutual funds in this research. Consequently, this paper excludes some mutual funds from the data set, which set the total of US open-end mutual funds for this part of the research to 668, split in 309 US open-end mutual funds with an one star score in terms of the Morningstar Sustainalytics rating, and 359 US open-end mutual funds with a five star score in terms of the Morningstar Sustainalytics rating. Here, this paper obtains TNA and returns on a monthly basis, and the interception date for all the 668 mutual funds. Table 6 in the Appendix shows the descriptive statistics of the fund characteristics.

4 Methodology

In order to investigate whether high rated US open-end mutual funds outperform low rated US open-end mutual funds, this paper applies several performance measurements to these different sets of mutual funds. First, this paper compares the performance of 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating, 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and all the 724 US open-end mutual funds over the period January 2000 - February 2021. To measure the performance of the funds, Section 4.1 conducts some of the performance measurement models of Eling and Faust (2010). These performance measurement models involve the classical single-factor (1) Capital Asset Pricing Model (CAPM), which is further extended to the inclusion of the factors introduced by (2) Fama and French (1993) and (3) Carhart (1997). Then, the performance of the funds is measured by the adjusted R^2 and the alpha distribution of the models.

Second, this paper examines different performance statistics, to investigate the performance of the different sets of US open-end mutual funds once again, together with the performance of the factors (Market proxy, SMB, HML and MOM), as passive benchmark indices. Section 4.3 defines the performance statistics in combination with formula. Then, the performance statistics of the US open-end mutual funds and the passive benchmark indices are ranked by performance.

Finally, this paper observes the relative flows of the US-open end funds. Here, the data set of Section 3.1 is applied. Therefore, the data set here ranges from March 2018 - February 2021 of 668 US open-end mutual funds, split in 309 US open-end mutual funds with an one star Morningstar Sustainalytics rating, and 359 US open-end mutual funds with a five star Morningstar Sustainalytics rating. Section 4.4 describes the regression and its parameters in further detail.

4.1 Performance measurement models

This section covers the introduction and description of the performance models of Eling and Faust (2010). First, an equally weighted portfolio is constructed using the different sets of US open-end mutual funds. For instance, for the high rated mutual funds, the equally weighted portfolio is constructed by taking the weighted average of the returns of all the high rated mutual funds at time t, for $t = 1, ..., T$. This delivers a vector of length T of weighted returns for the three sets of mutual funds, the same length as the individual fund returns.

To properly compare the results, the adjusted R^2 is compared for the different sets of mutual funds using the different performance models. The Adjusted R^2 of a model is defined as

$$Adjusted \quad R^2 = 1 - \frac{(1 - R^2)(N - 1)}{N - p - 1}, \tag{1}$$

with N is the number of observations (mutual funds), p is the number of independent variables in the model and R^2 is the sample R-squared, which is a statistical measurement of the portion of the variance that a linear model explains, by dividing the portion of variance that the model explains by the total variance. Table 2 in the Appendix shows the adjusted R^2 of the models using the equally weighted portfolios, and the distribution of the adjusted R^2 for the individual funds.

The constant term of the performance models (α) could function as a performance measurement as well. For instance, a negative constant could indicate that the mutual fund manager underperforms relative to a passive benchmark. Therefore, to test the first hypothesis whether sustainable mutual fund managers more often outperform their relative benchmark than their non-sustainable peers, this paper looks at the alpha distribution among the different funds, and test for a potential significant outperformance for the different performance measurement models. It does so, by looking at the z-statistic

$$z = \frac{\hat{p}_L - \hat{p}_H}{SE_P}, \tag{2}$$

where $\hat{p}_L$ and $\hat{p}_H$ are the percentages of alphas that are larger than zero for the individual low rated funds and high rated funds, respectively. The standard deviation for the pooled percentage of positive alphas (SE_P) is given by

$$SE_{DP} = \sqrt{\hat{p}(1 - \hat{p})(\frac{1}{n_L} + \frac{1}{n_H})}, \tag{3}$$

where $\hat{p}$ is the pooled percentage of alphas larger than zero and n_L and n_H are the number of low rated funds and high rated funds, respectively. Table 3 in the Appendix shows the constant terms of the models using the equally weighted portfolios, the distribution of the constant terms for the individual funds, and the corresponding z-statistics for the different performance measurement models.

4.1.1 Capital Asset Pricing Model

The first performance measurement model is the Capital Asset Pricing Model (CAPM). This model is based on the modern portfolio theory of Markowitz (1952), and is defined as

$$R_{it} - R_{ft} = \alpha_i + \beta_i(R_{mt} - R_{ft}) + \epsilon_{it}, \tag{4}$$

where R_{it} is the return of fund i at time t, R_{ft} is the risk-free return at time t, α_i is the intercept of the regression, β_i is the beta of fund i, R_{mt} is the market return at time t, and ϵ_{it} is the error term of fund i at time t. The intercept of this regression is a risk-adjusted performance measurement on its own, introduced by Jensen (1968). In this case, Jensen's alpha determines the abnormal return of fund i relative to the theoretical return expected by the CAPM model.

4.1.2 Fama and French (1993) three-factor model

The next model is a well-known extension on the CAPM model of Section 4.1.1. Fama and French (1993) introduce two extra factors in this model. They consider the term $(R_{mt} - R_{ft})$ as a market factor that describes the excess return of the market relative to the risk-free return. Therefore, they add two extra factors, SMB and HML, to obtain a three-factor model. The Small Minus Big (SMB) factor describes the expected outperformance of small companies relative to big companies. The High Minus Low (HML) factor looks at the market-to-book ratio of companies and described the expected outperformance of companies with a high market-to-book ratio relative to companies with a low market-to-book ratio. This is also called the value factor, since the factor says that value companies outperform growth companies. This three-factor model, together with the Carhart (1997) model of Section 4.1.3, are considered to be the best models to measure the performance of mutual funds (Otten and Bams (2002)). The Fama and French (1993) three-factor model is defined as

$$R_{it} - R_{ft} = \alpha_i + \beta_{i,m}(R_{mt} - R_{ft}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \epsilon_{it}, \tag{5}$$

where $\beta_{i,p}$ is the beta of fund i for factor p with $p = \{m, SMB, HML\}$, SMB is the Small Minus Big factor, and HML is the High Minus Low factor on time t.

4.1.3 Carhart (1997) model

The Carhart (1997) model adds another factor to the model. The Monthly Momentum (MOM) factor tries to capture the momentum effect of the funds. To do so, it looks at if the prior 12-month average of returns is positive. If this is the case, the stock would consider to show momentum. Adding the MOM factor to the model, the Carhart (1997) model is defined as

$$R_{it} - R_{ft} = \alpha_i + \beta_{im}(R_{mt} - R_{ft}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,MOM}MOM_t + \epsilon_{it}, \tag{6}$$

where $\beta_{i,s}$ is the beta of fund i for factor s with $s = \{m, SMB, HML, MOM\}$, and MOM is the Monthly Momentum factor on time t.

4.2 Sub periods

The alphas of the models of Section 4.1 are constructed via multiple data sets. However, the models all use the same time period. This could influence the results. For instance, if there is a financial crisis during the entire time period, results could change dramatically. Therefore, to investigate whether the performance models are robust for times of financial distress during the sample period, this research split the entire time period into five different subperiods of Eurola (2022). Each of these subperiods represent a different time of financial distress. The subperiods are:

- ICT bubble: 04/2000 - 08/2002,
- Economic boom: 09/2002 - 10/2007,
- Financial crisis: 11/2007 - 02/2009,
- Recovery periods: 03/2009 - 01-2020, and
- Covid-19 period: 02/2020 - 02/2021.

Table 4 in the Appendix shows the same results as Table 3, only this time for the five different subperiods.

4.3 Performance measurement statistics

This section introduces different types of performance measurement statistics, their relative ranking and their test statistics for the equally weighted portfolios, as in Section 4.1, and the market proxy, and the factors SMB, HML, and MOM, as passive benchmark indices. Table 4.3 in the Appendix shows the performance measurement statistics and test statistics for the portfolios and indices.

4.3.1 Sharpe ratio

A statistical measurement for the risk (volatility) of an investment against the expected excess return is the Sharpe ratio of Sharpe (1994). The Sharpe ratio is given by

$$SR = \frac{R_i - R_f}{\sigma_i}, \tag{7}$$

where R_i is the return of fund i, R_f is the risk-free return, and σ_i is the standard deviation of the excess return of fund i.

To test the second hypothesis, this paper test for potential significant differences between Sharpe ratios among the different set of funds. It does so by applying the Jobson and Korkie (1981) test. Jobson and Korkie (1981) introduce a test that compares Sharpe ratios and investigates a potential significant different between these Sharpe ratios. The $z$statistic for the Jobson and Korkie (1981) test is given by

$$\hat{z}_{JK} = \frac{\hat{\sigma}_j\hat{\mu}_i - \hat{\sigma}_i\hat{\mu}_j}{\sqrt{\hat{\phi}}}, \tag{8}$$

where $\hat{\mu}_i$ ($\hat{\mu}_j$) is the mean of fund i (fund j), $\hat{\sigma}_i$ ($\hat{\sigma}_j$) is the standard deviation of fund i (fund j), and $\hat{\phi}$ equal to

$$\hat{\phi} = \frac{1}{T}(2\hat{\sigma}_i^2\hat{\sigma}_i^2 - 2\hat{\sigma}_i\hat{\sigma}_j\hat{\sigma}_{i,j} + \frac{1}{2}\hat{\mu}_i^2\hat{\sigma}_j^2 + \frac{1}{2}\hat{\mu}_j^2\hat{\sigma}_i^2 - \frac{\hat{\mu}_i\hat{\mu}_j}{\hat{\sigma}_i\hat{\sigma}_j}\hat{\sigma}_{i,j}^2), \tag{9}$$

with $\hat{\sigma}_{i,j}$ represents the covariance between fund i and fund j.

4.3.2 Modified Sharpe ratio (Israelsen (2003) and Pézier, White, et al. (2006))

The modified Sharpe ratio is a similar statistical measurement as the Sharpe ratio in Section 4.3.1, where the Israelsen (2005) slightly adjust the denominator. Therefore, the modified Sharpe ratio of Israelsen (2005) is given by

$$MSR = \frac{ER}{\sigma_i^{\frac{ER}{|ER|}}}, \tag{10}$$

where $ER = R_i - R_f$ and $|ER|$ is the absolute value of $R_i - R_f$. So, instead of dividing the excess return of the security by the standard deviation, it divides the excess return by the standard deviation to the power of a extra term. This means that the denominator of the Sharpe ratio increases in value when the excess return of the security is positive, and decreases in value when the excess return of the security is negative.

The adjusted Sharpe ratio of Pézier, White, et al. (2006) uses the Sharpe ratio of Section 4.3.1 in their formula. Furthermore, they also use the third moment (skewness) and fourth moment (kurtosis) into their calculation. Therefore, the adjusted Sharpe ratio of Pézier, White, et al. (2006) is given by

$$ASR = SR\left[1 + \frac{S}{6}SR - \left(\frac{(K-3)}{24}\right)SR^2\right], \tag{11}$$

where SR is the Sharpe ratio of 7, S is the skewness and K is the kurtosis. When the skewness and kurtosis follow a normal distribution, they are equal to zero and three, respectively. This means that, if the skewness and kurtosis follow a normal distribution, the adjusted Sharpe ratio of Pézier, White, et al. (2006) is equal to the Sharpe ratio of Section 4.3.1. When the skewness and kurtosis are not equal to zero and three, the adjusted Sharpe ratio corrects for this non-normal distribution.

4.3.3 Sortino ratio

The Sortino ratio is another differentiation of the Sharpe ratio of Section 4.3.1. Again, it only adjusts the denominator of the formula. This time, a distinction is made between different types of volatility. Whereas the Sharpe ratio in Section 4.3.1 uses the entire standard deviation, the Sortino ratio only uses the standard deviation of the returns that fall below a given threshold value, also referred to as the Minimum Acceptable Return (MAR). This is called the downside deviation. Usually, the MAR is equal to zero. This means that the Sortino ratio only uses the stardard deviation of the negative returns. Therefore, whereas the Sharpe ratio is a statistical measurement for the risk of an investment against the expected excess return, the Sortino ratio is a statistical measurement for the bad risk of an investment against the expected excess return. The formula for the Sortino ratio is equal to:

$$Sortino = \frac{R_i - R_f}{\sigma_d}, \tag{12}$$

where σ_d is the downside deviation, which is a measurement of the risk that the return of fund i falls below a given threshold value. In this case, when the return of fund i falls below zero.

4.3.4 Treynor ratio

The Treynor ratio is again a statistical performance measurement that investigates the relationship between the excess return of an investment against a risk factor. Here, the nominator is again the excess return of the security above the risk-free rate. The denominator is a different measurement for risk, in particular, systematic risk. This is measured as the beta of the fund. Therefore, the Treynor ratio is given by

$$Treynor = \frac{R_i - R_f}{\beta_i}, \tag{13}$$

where β_i is defined as the beta of fund i, and is defined as

$$\beta_i = \frac{Cov(R_i, R_f)}{Var(R_f)}, \tag{14}$$

where $Cov(R_i, R_f)$ is equal to the covariance of the return of fund i and the risk-free rate, and $Var(R_f)$ is equal to the variance of the risk-free rate. The beta of a fund indicates the extent to which the fund moves with shocks in the market. For example, a beta equal to one means that the fund moves exactly like the market.

4.4 Mutual fund flows

Another way to get an indication about the performance of mutual funds, is to look at the fund flows. In particular, to look at the net relative flows of a mutual fund. This paper applies the formula for net relative flows of Gruber (1996). To determine the net relative flows, the net absolute flows should be calculated first. These are given by

$$F_{i,t} = TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}), \tag{15}$$

where $TNA_{i,t}$ denotes the Total Net Assets (TNA) of fund i at the end of month t and $R_{i,t}$ is return of fund i at the end of month t. Then, the formula for the net relative flows is equal to

$$f_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t})}{TNA_{i,t-1}} = \frac{F_{i,t}}{TNA_{i,t-1}}. \tag{16}$$

A typical regression model that specifies the net relative flows as a linear function includes some past performance measure and a set of control variables (Goriaev, Nijman, Werker, et al. (2002)). In this case, as the performance measure, this paper chooses the Jensen's alpha for fund i. To calculate the alphas, this paper constructs a rolling window of six months. For example, the first Jensen's alpha for fund i is determined with the first six months of the data set. Since the data set for the Total Net Assets ranges from March 2018 - February 2021 (36 months), this paper obtains 30 Jensen's alpha for every fund. This paper also includes the lagged relative flow of fund i in the regression, since it seems logic that there is a strong relationship between the relative flows at time t and at time $t - 1$. To observe the difference between the low rated funds and the high rated funds, this paper introduces a dummy variable, that is given by

$$D_{i,t} = \begin{cases} 1, & \text{if fund } i \text{ is high rated,} \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

The dummy variable is both individually and attached to the lagged relative flow and the Jensen's alpha included in the regression. to properly obtain the differences between the different set of mutual funds.

For the set of control variances, this paper chooses the same set as the paper of Goriaev, Nijman, Werker, et al. (2002), which contains the fund characteristics size and age. It seems to restrictive to assume that the regression coefficients do not depend on any fund characteristic. Therefore, the regression model looks like

$$\begin{aligned} f_{i,t} =& \gamma_0 + \gamma_1 f_{i,t-1} + \gamma_2 D_{i,t} + \gamma_3 D_{i,t} x f_{i,t-1} + \gamma_4 \alpha_{i,t} \\ &+ \gamma_5 D_{i,t} x \alpha_{i,t} + \gamma_6 log(TNA_{i,t-1}) + \gamma_7 log(Age_{i,t-1}) + \epsilon_{i,t}, \qquad t = 6, 7, ..., T \end{aligned} \tag{18}$$

where γ_0 represents the constant term, $\alpha_{i,t}$ represents Jensen's alpha for fund i at time t, $log(TNA_{i,t-1})$ and $log(Age_{i,t-1})$ represents the natural logarithms of the size and age characteristic, respectively. Taking the logarithm of size and age improves the fit of the model. This model deals with panel data, a data set over funds and over time, this model requires two stage Fama and MacBeth (1973) regression. This means that, in the first stage, this model applies time-series regression, and, in the second stage, a cross-sectional regression. Eventually, this leads to single estimators for the eight coefficients. Table 4.4 in the Appendix shows the coefficients of the panel regression and the corresponding standard errors.

5 Results

5.1 Descriptive statistics

Table 1 shows the descriptive statistics for the US open-end mutual funds returns and for the returns of the four risk factors. The means of the different mutual fund selections do not significantly differ from each other. The same holds for the standard deviations. However, the means are substantially higher than the ones of the factors, especially for HML. However, in terms of standard deviations, the ones for the factors are lower than the standard deviations of the mutual funds. The one exception is the standard deviation for the market proxy, which is even larger than the standard deviations for the mutual funds.

Whereas some investors care more about means and standard deviations, others care more about skewness and kurtosis. Table 1 reports a negative skewness and a positive kurtosis for all funds. However, especially for the kurtosis, the low rated mutual funds report more extreme values. Incontrast to the skewness of the funds, the skewnesses of the factors are all positive, except for the market proxy. In terms of kurtosis, only the momentum factors report a negative value, whereas the kurtosis for the size factor is most positive. As Scott and Horvath (1980) states in his research, most risk averse investors seek higher values with odd numbers (mean and skewness), and lower values with even number (standard deviation and kurtosis). Therefore, these types of investors could favor the high rated mutual funds, although the differences are not significant. Therefore, because the differences do not seem significant, investors could "do well" and still perform the same, in contrary to the findings of Statman and Glushkov (2009), where they find evidence for this hypothesis.

5.2 Performance measurement models

5.2.1 Adjusted R^2

Table 2 shows the adjusted R^2 for equally weighted portfolios of the different mutual funds sets, estimated with (1) the Capital Asset Pricing Model, (2) the Fama and French (1993), and (3) the Carhart (1997) model over the period January 2000 - February 2021. Moreover, Table 2 also shows the minimum and maximum adjusted R^2, the 25%, and 75% quantile and the median adjusted R^2 for all the individual mutual funds.

For the equally weighted portfolios, Table 2 reports the lowest adjusted R^2 for the CAPM models. For the other models, there is no significant difference in adjusted R^2 for all the portfolios. This means that the explanatory power of the model increases when the size and value factors are included in the model, but do not increase when the momentum factor is included as well. This means that the momentum factor has no relationship with the excess returns of the portfolios. In terms of the equally weighted portfolios, the models using the high rated portfolio explain the least amount of variation in the mutual fund returns, although the differences are not substantially large again. For example, for the Fama and French model, using the low rated portfolio explain about 96.8% of the valuation of mutual fund returns, whereas the model only explains about 95.3% of the valuation of returns using the high rated portfolio. All the adjusted R^2 for the equally weighted portfolio of all funds are higher than for the other portfolios. One argument may be that this portfolio is a combination of more mutual funds, which makes it a more diversified portfolio. On average, more diversified portfolios perform better than their less diversified peers.

In terms of the individual funds, Table 2 reports similar values for the different mutual fund selections. However, the minimum adjusted R^2 for the CAPM model for the high rated funds and for the entire set of funds are negative, whereas the adjusted R^2 for the CAPM model for the low rated funds is positive. Therefore, individual funds that obtain a negative adjusted R^2, obtain insignificant the explanatory variables. All the other values do not substantially differ from each other. However, Table 2 shows that, on average, the explanatory power is lower for the individual funds than for the equally weighted portfolios. In particular, the medians for the individual funds are lower than the adjusted R^2 for the equally weighted portfolios. For example, the median adjusted R^2 for the low rated funds for the Carhart model gives an explanatory power of 84.9%, where the equally weighted portfolio using this model reports a value of 96.8%. Again, this could be the case because the equally weighted portfolios are more diversified, which contributes to performance. Finally, the values for the Fama and French model and the Carhart model for the individual funds do not differ significantly, again pointing out that including the momentum factor seems of no purpose.

This means that, looking at the different adjusted R^2, there is no significant different in performance in terms of a more sustainable oriented portfolio against its less sustainable peer. Moreover, the portfolio containing all the funds seems to be in favor, however, this could only be the case due to the increase in the number of funds included. In terms of performance models, the Fama and French model and the Carhart model obtain the most explanatory power, where including the momentum factor seems negligible.

5.2.2 Alpha distribution

Table 3 shows the alphas for equally weighted portfolios of the different mutual funds classes, estimated with (1) the Capital Asset Pricing Model, (2) the Fama and French (1993), and (3) the Carhart (1997) model over the period January 2000 - February 2021. Moreover, Table 3 also shows the minimum and maximum adjusted alpha, the 25%, and 75% quantile and the median adjusted alpha for all the individual mutual funds. Finally, Table 3 shows the distribution of the significant negative and positive alphas.

For the equally weighted portfolios, most values for alpha are positive, except for the alphas of the Carhart model. This indicates that mutual fund managers underperform passive benchmarks according to this model. However, all these alphas for the equally weighted portfolio are not significantly different from zero. In terms of comparing the low rated portfolio against the high rated portfolio, only the alphas from the Fama and

French model seem substantially different between the two portfolios. As in Section 5.2.1, Table 3 shows results that may indicate that adding the momentum factor to the model seems negligible. However, in this case, it seems useless to add the size and value factor as well, since the CAPM model obtains the highest alphas. Nevertheless, all alphas are insignificant.

The individual funds report higher average alphas than the equally weighted portfolios. For example, the median alpha for high rated funds using the CAPM model is equal to 0.717, whereas the alpha for a high rated portfolio using the CAPM model is equal to 0.070. Next, all the minimum values of the alphas for the individual funds are negative, suggesting that there are mutual fund managers that underperform passive benchmarks, for both sustainable mutual funds and non-sustainable mutual funds. Again, the alphas of the Carhart model are most negative. In terms of alpha distribution, the percentage of mutual fund managers that outperform a passive benchmark (sign.>0) is higher for high rated funds than for low rated funds. For instance, using the Fama and French model, high rated mutual fund managers outperform the benchmark in 20.66% of the cases, relative to 13.55% of their peers. This indicates that high rated mutual fund managers perform better than their low rated rivals.

To test whether the high rated mutual fund managers significantly have a higher outperformance relative to low rated fund managers (H_1), this paper constructs a test statistic for the possible significant difference between the percentages of alphas of the individual funds that outperform their passive benchmark (sign.>0). Table 3 shows that only for the Fama and French model, the high rated fund managers outperform low rated fund managers, on a 5% significance level. Therefore, using the Fama and French model, this paper do not reject H_1, since there is significant evidence that the high rated fund managers more often outperform their passive benchmark relative to low rated fund managers. For the other two performance measurement models, this paper finds no significant evidence for this outperformance. Therefore, in these cases, this paper reject H_1. There is no significant evidence that high rated funds more often outperform their passive benchmark relative to low rated funds.

5.2.3 Robustness check

Table 4 reports alphas for equally weighted portfolio, estimated with (1) the Capital Asset Pricing Model, (2) the Fama and French (1993), and (3) the Carhart (1997) model over the sub periods April 2000 - August 2002, September 2002 - October 2007, November 2007 - February 2009, March 2009 - January 2020, and February 2020 - February 2021. Moreover, Table 4 also shows the minimum and maximum adjusted alpha, the 25%, and 75% quantile and the median adjusted alpha for all the individual mutual funds. Finally, Table 4 shows the distribution of the significant negative and positive alphas.

Table 3 shows, for the equally weighted portfolios, all positive values for alpha, except for the Carhart model. This pattern is most comparable with the subperiods April 2000 - August 2002 and November 2007 - February 2009. On average, the other subperiods show more significant values for alpha, except for the last subperiod. The first subperiod shows only significant positive values for the alphas of the CAPM model. During this subperiod, there is no significant preference for the different portfolios. The low rated portfolio performs best using the CAPM model, while including the size and value factor in the model put the high rated portfolio in the top spot. Similar to Table 3, adding the momentum effect seems negligible. The subperiod September 2002 - October 2007 shows only positive values, with the high rated portfolio as the best performer. The subperiod November 2007 - February 2009 gives similar results as the first subperiod. The main differences are that the CAPM alphas are not significant here, and that the high rated portfolio comes out as the best performer. The subperiod March 2009 - January 2020 shows only significant negative values, completely opposite to the results of the subperiod September 2002 - October 2007. The final subperiod shows only positive values for alpha, except for the alpha of the low rated portfolio using the CAPM model. Here, all values for alpha are not significant. Splitting the research periods into five different subperiods offer more significant values for the alphas, however, it gives different sign patterns for the alphas. Overall, Table 4 finds evidence that the alphas for the equally weighted portfolios are not robust for different subperiods.

In terms of the alphas for the individual portfolios, on average, most individual funds report lower alphas than the equally weighted portfolio during the different subperiods. This is different from the individual funds over the entire sample period, where they report higher values for alpha than the equally weighted portfolios. Again, this could mean that the performance models are not robust for different sample periods. The two main similarities with these results and the results of Table 3 are that all minimum values for the individual funds are negative, and that, in most cases, the Carhart model obtains the most negative values for alpha. This again could mean that there are fund managers that underperform relative to the benchmark and that the momentum factor is negligible in the model. In terms of alpha distribution, the percentage of mutual fund manager that outperform a passive benchmark (sign.>0) is higher for high rated funds than for low rated funds in the periods November 2007 - February 2009 and March 2009 - January 2020, but lower in the other subperiods. Therefore, there is no unanimous indication about the performance of high rated mutual funds managers relative to their low rated peers during the subperiods. This means that there is no significant indication to accept the hypothesis H_1.

5.3 Performance measurement statistics

Table 5 reports performance measurement values and their rankings for the open-end mutual funds, the market proxy, the Small minus Big (SMB) factor, the High minus Low (HML) factor, and the Momentum (MOM) factor over the period January 2000 - February 2021.

Table 5 shows that the high rated funds obtain higher Sharpe ratios than the other funds, with only the momentum factor obtaining higher values. This means that the average risk-adjusted return for the high rated funds are higher than the risk-adjusted returns of their peers. In terms of the Sortino ratio, the high rated funds are performing the worst, with the low rated funds ending on top. This means that low rated funds have a lower downside deviation than high rated funds. In particular, the standard deviation of the negative returns for low rated funds are substantially lower than the standard deviation of the negative returns for high rated funds. Since the Sharpe ratio is higher for the high rated funds, and the Sharpe ratio uses the entire standard deviation, the opposite relationship could be true for the standard deviations of the positive returns for the low rated funds and high rated funds. Finally, the Treynor ratio of different sets of funds are all lower than those of the factors. The Treynor ratio of the high rated funds is by far the most negative, with no substantial differences between the Treynor ratio of the low rated funds and all the funds. This could indicate that the systematic risk factor (beta) for all the sets of funds are negative. Since the Treynor ratio of the high rated funds is by far the most negative, the systematic risk factor could be smaller negative than for the other sets of funds. This means that all the sets of funds react to shocks in the market in the opposite way than the market, where the high rated funds react the least.

To test whether the high rated funds statistically outperform the low rated funds (H_2), this paper applies the Jobson and Korkie (1981) test, to test whether the Sharpe ratios of the different sets of funds significantly differ from each other. Table 5 shows test statistics for the Jobson and Korkie (1981) test, and finds no significant different between the Sharpe ratios. Therefore, this paper rejects H_2, since there is no evidence that the Sharpe ratio for the high rated funds significantly outperform the Sharpe ratio for the low rated funds.

5.4 Fund characteristics

Table 6 reports means and corresponding standard deviations of the flows, performance size and age for the 309 US open-end mutual funds with an one star Morningstar Sustainalytics rating, the 359 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and all the 668 US open-end mutual funds over the period March 2018 - February 2021.

Table 6 shows negative average absolute flows for all the different sets of mutual funds. This indicates that, on average, the funds in the sample have larger outflows than inflows. The average absolute flow for the high rated funds is less negative than the average absolute flow for the low rated funds, indicating better performance for high rated funds. However, the standard deviations for the different sets of funds are significantly larger than the absolute value of the corresponding means. Therefore, the negative average absolute flows could be the cause of some outliers in the sample. The means for the relative flows are larger than zero for the high rated funds and the full sample of funds, but negative for the low rated funds. Relative flows are defined as the percentage growth of assets of the fund. Therefore, an average negative relative flow indicate that assets of the low rated funds, on average, have a decrease in value. The average relative flow for the full sample becomes positive due to the high rated portion of funds. Similar to the absolute flows, the standard deviation for the relative flows are significantly larger than the means, possibly caused by outliers.

The average Jensen's alpha for the low rated funds is negative. This indicates an average negative abnormal return of the individual funds. Therefore, the average low rated fund have a underperformance relative to their passive benchmark. However, the mean for the alphas is not significantly lower than zero, due to the standard deviation of 1.645. Contrary to the low rated funds, the average Jensen's alpha for the high rated funds is positive, suggesting a positive abnormal return for the individual high rated funds. Consequently, the high rated funds, on average, outperform their passive benchmarks. Nevertheless, the mean is not significantly different from zero, indicating that the individual high rated funds, on average, perform similarly than their passive benchmarks. The full sample average is also negative, due to the negative average performance of the low rated funds. Therefore, the funds in the full sample, on average, underperform their relative benchmark. However, they do not significantly underperform their benchmarks.

Table 6 also shows the average size (Total Net Assets) and age and the corresponding standard deviations for the different sets of funds. This paper chooses size and age as characteristics in regression of Equation 18. The table shows a larger average size for the low rated funds, but a larger average age for the high rated funds. As stated in Section 2, there are money shifts from non-sustainable investments to sustainable investments, However, as Table 6 shows, the low rated funds are still larger than the high rated funds. Therefore, the money shift does have to increase further. Since the standard deviation for the size of the low rated funds is significantly larger than the standard deviation for the high rated funds, the set of low rated funds is more divided between small-cap funds and large-cap funds than the set of high rated funds. According to these findings, where the average size for the low rated funds is larger than for the high rated funds, and the average performance is lower for the low rated funds than for the high rated funds, fund size could harm performance. This in line with existing literature (Chen et al. (2004)). The average age for the sets of funds is significantly different from each other. Whereas the low rated funds have an average age of around 2 years, the high rated funds have an average age of around 18 years. However, this does not mean that all the high rated funds are focused on sustainability since their inception date. It could be that the individual funds started focusing on sustainability since a couple of years, while they began investing in the market far before that period. On average, investors prefer funds with longer operating histories, since this give them more insights in the performance of the fund.

5.5 Mutual fund flows

Table 7 reports coefficient estimates of the panel regression of Equation 18 over the period March 2018 - February 2021. The dependent variable is fund net relative flow. The explanatory variables include a constant, the lagged relative flow, a dummy variable, the log of fund size, the log of fund age, and Jensen's alpha, estimated over a period of six months.

Table 7 shows a negative coefficient for the dummy variable ($D_{i,t}$). This indicates that the high rated funds have lower relative flows than the low rated funds. However, this coefficient is not significantly different from zero. This means that there is no significant difference between the relative flows of the low rated funds and the high rated funds. Table 6 shows the same result. So, again, there is a money shift towards sustainable investing in the financial market, but this shift is significant in the data set. Section 2.1.1 state that the critical point for Corporate Social Responsibility (CSR) lies around 2012 (Visser (2011)), and the data for this regression starts in 2018, there is still no significant money shift to sustainable mutual funds. Therefore, although there is a globally emphasis to battle global warming and investing in corporations that support this, there still is a need for the shift to increase in the upcoming years.

Table 7 shows a negative relationship between the relative flow at time t and the lagged relative flow at time $t-1$. This indicates that there is a negative momentum effect in the relative flows of the individual funds. Therefore, the mutual funds flows move in the opposite direction, indicating a negative growth value for the fund flows at time t when the fund flows grew at time $t-1$, and vice versa. However, the coefficient for the lagged relative flow is not significantly different from zero. This means that the lagged value does not significantly determine the level of the current relative flow. This is in contrary to intuition. Moreover, it seems logic that the lagged relative fund flows have an extensive positive relationship with the relative fund flows at time t. Moreover, Table 7 shows a coefficient for $D_{i,t}$ x $f_{i,t-1}$ of -0.001. This also indicates no significant relationship between the relative flow at time t and the lagged relative flow at time $t-1$, specific for the high rated funds.

Furthermore, Table 7 shows the coefficients and standard deviation for the size and age characteristics. The coefficients for the size and age characteristic are both negative. This means that the regression results suggest a negative relationship between both the size of the fund and its relative flows, and the age of the fund and its relative flows. In particular, smaller and younger funds find, on average, more relative growth in their assets. These findings are in line with existing literature, since Chevalier and Ellison (1997) also state that smaller and younger funds have larger relative flows than their peers. However, similar to the previous regression results, these coefficients are both not significantly different from zero. Therefore, these results should be interpreted with caution, since there is no significant relationship between the size and age of the fund, and the corresponding relative flows.

The positive coefficient for the performance variable suggest a positive relationship between the relative flows and the financial performance, in the form of the Jensen's alpha. Therefore, the relative fund flows are sensitive to past performance. If the individual fund performs well during the month, it positively effects the growth of its net assets. Additionally, attaching the dummy variable to the performance gives a negative coefficients. This suggests that the performance of high rated funds have a negative effect on their relative flows. Moreover, for the low rated funds, when the dummy variable is set to zero, the coefficient for the performance variable is equal to 0.396, and, for the high rated funds, when the dummy variable is set to one, the coefficients is equal to $0.396 - 0.558 = -0.162$. This means that a good performing low rated fund increases its relative flows, whereas the flows for high rated funds increases in bad performing times. However, both the coefficients are not significantly different from zero, mainly due to the high standard errors of these coefficients. Therefore, both for low rated funds and high rated funds, the individual performance does not have a significant effect on its relative flows. Thus, since the coefficient for the performance of high rated funds ($D_{i,t}$ x $\alpha_{i,t}$) is not coefficient, this paper reject the hypothesis that high rated fund flows are less sensitive to past performance than low rated fund flows (H_3).

6 Conclusion

This paper investigates whether sustainable oriented mutual funds outperform their non-sustainable peers. Therefore, the research question here is whether US open-end mutual funds that construct a sustainable portfolio outperform their less sustainable peers. First, to answer this research question, this paper looks at the potential outperformance of high rated US open-end mutual funds and low rated US open-end mutual funds relative to their respective benchmarks. Second, this paper compares multiple performance statistics of the two different sets of mutual funds. Finally, this paper conducts a regression to find a possible relationship about the relative flows of the mutual funds.

The potential outperformance is measured via several performance measurement models. In particular, the Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model, and the Carhart (1997) model. This paper looks at the adjusted R^2 and the alpha distribution of the models using the different sets of mutual funds. The adjusted-R^2 does not indicate a significant preference for either of the mutual fund sets. The set containing all the mutual funds obtain a slight advantage, however, this could be due to a higher diversification rate. In terms of models, the Fama and French model and the Carhart model obtain the most explanatory power, where including the momentum factor in the model seems negligible.

Almost all alphas are returning positive values for the equally weighted portfolios, except for the alphas of the Carhart model. This means that the Carhart model obtains values for alpha where the mutual funds underperform their relative benchmark. These values for alpha should be interpreted with caution, since none of the alphas are significantly different from zero. Furthermore, the individual funds obtain a higher alpha than the equally weighted portfolios. The percentage of mutual fund managers that outperform their passive benchmark is higher for high rated funds than for low rated funds, indicating that high rated mutual funds outperform their low rated peers. However, the outperformance is not significant. Moreover, the alpha distribution does not seem robust for different sets of data as well. Dividing the total sample period into five different sub periods as a robustness check for the alpha distribution, gives different results than using the full sample. Therefore, the first hypothesis that high rated funds more often outperform their passive benchmark relative to low rated funds (H_1) is rejected.

Similar to the results of the adjusted R^2, the performance measurement statistics do not prefer one mutual funds set over another. There are some statistics in favor of the high rated mutual funds, and there are statistics preferring the low rated mutual funds. Therefore, based on the performance measurement statistics, there is no significant preference between the set of more sustainable mutual funds over their less sustainable peers. Moreover, the Sharpe ratios for the low rated funds and the high rated funds do not significantly differ from each other. Thus, the second hypothesis that high rated funds outperform low rated funds based on performance statistics (H_2) is rejected.

The regressions containing the relative flows of the mutual funds all obtain insignificant coefficients for the performance variable as well as for the other explanatory variables in relationship to the relative flows as dependent variable. The coefficient for the performance variable in combination with the dummy variable is insignificant as well. Therefore, there is no significant indication that high rated fund flows are less sensitive to performance than low rated fund flows. Thus, the final hypothesis that high rated fund flows are less sensitive to performance than low rated fund flows (H_3) is rejected.

Overall, there seems no significant difference between the performance of the different sets of mutual funds. Since this indicates that this paper rejects all the hypotheses, this paper concludes that the US open-end mutual funds that construct a sustainable portfolio do not significantly outperform their less sustainable peers. However, to end with a positive note, they do not significantly underperform as well. This could still have a positive effect on the financial market. Moreover, when the financial performance of investors do not decrease when investing in sustainable securities, investors could focus more on sustainable investing without compromising on their financial performance. Eventually, this shift could increase financial performance as well. Thus, investors sure can "do good and do well".

Nevertheless, there are some limitations available in this report. First, the data for the mutual funds flows is limited in this paper. Therefore, the results may differ from previous literature. Moreover, the entire data set only comes from the United States. Although the United States is a large and important part of the financial market, it does not represent the entire financial market. These data limitations could influence the total research. Therefore, for further research, one could focus on European funds or take sets of mutual funds from different parts of the world.

References

[1] Nerilie J Abram, Helen V McGregor, Jessica E Tierney, Michael N Evans, Nicholas P McKay, and Darrell S Kaufman. "Early onset of industrial-era warming across the oceans and continents". In: *Nature* 536.7617 (2016), pp. 411–418.

[2] Alicia Adamczyk. "Millennials spurred growth in sustainable investing for years. Now, all generations are interested in ESG options". In: (2021). URL: https://www.cnbc.com/2021/05/21/millennials-spurred-growth-in-esg-investing-now-all-ages-are-on-board.html.

[3] Timothy Adler and Mark Kritzman. "The cost of socially responsible investing". In: *The Journal of Portfolio Management* 35.1 (2008), pp. 52–56.

[4] George A Akerlof. "A theory of social custom, of which unemployment may be one consequence". In: *The quarterly journal of economics* 94.4 (1980), pp. 749–775.

[5] Global Sustainable Investment Alliance. *Global sustainable investment review.* 2018.

[6] Manuel Ammann, Christopher Bauer, Sebastian Fischer, Philipp Mueller, et al. *Tha Impact of the Morningstar Sustainability Rating on Mutual Fund Flows.* Tech. rep. University of St. Gallen, School of Finance, 2017.

[7] Manuel Ammann, Christopher Bauer, Sebastian Fischer, and Philipp Müller. "The impact of the Morningstar Sustainability Rating on mutual fund flows". In: *European Financial Management* 25.3 (2019), pp. 520–553.

[8] Doron Avramov and Russ Wermers. "Investing in mutual funds when returns are predictable". In: *Journal of Financial Economics* 81.2 (2006), pp. 339–377.

[9] Gary S Becker. *The economics of discrimination.* University of Chicago press, 2010.

[10] Terry Beckman, Alison Colwell, and Peggy H Cunningham. "The emergence of corporate social responsibility in Chile: The importance of authenticity and social networks". In: *Journal of Business Ethics* 86.2 (2009), pp. 191–206.

[11] Thomas C Berry and Joan C Junkus. "Socially responsible investing: An investor perspective". In: *Journal of business ethics* 112.4 (2013), pp. 707–720.

[12] Christopher R Blake and Matthew R Morey. "Morningstar ratings and mutual fund performance". In: *Journal of financial and Quantitative Analysis* 35.3 (2000), pp. 451–483.

[13] Richard T Bliss and Mark E Potter. "Mutual fund managers: does gender matter?" In: *The Journal of Business and Economic Studies* 8.1 (2002), p. 1.

[14] Marshall E Blume. "An anatomy of Morningstar ratings". In: *Financial Analysts Journal* 54.2 (1998), pp. 19–27.

[15] Shernaz Bodhanwala and Ruzbeh Bodhanwala. "Relationship between sustainable and responsible investing and returns: a global evidence". In: *Social Responsibility Journal* (2019).

[16] Nicolas PB Bollen. "Mutual fund attributes and investor behavior". In: *Journal of financial and quantitative analysis* 42.3 (2007), pp. 683–708.

[17] Stephen J Brown and William N Goetzmann. "Performance persistence". In: *The Journal of finance* 50.2 (1995), pp. 679–698.

[18] Mark M Carhart. "On persistence in mutual fund performance". In: *The Journal of finance* 52.1 (1997), pp. 57–82.

[19] Sudheer Chava. "Environmental externalities and cost of capital". In: *Management science* 60.9 (2014), pp. 2223–2247.

[20] Sudheer Chava. "Socially responsible investing and expected stock returns". In: *Available at SSRN 1678246* (2010).

[21] Joseph Chen, Harrison Hong, Ming Huang, and Jeffrey D Kubik. "Does fund size erode mutual fund performance? The role of liquidity and organization". In: *American Economic Review* 94.5 (2004), pp. 1276–1302.

[22] Judith Chevalier and Glenn Ellison. "Risk Taking by Mutual Funds as a Response to Incentives". In: *Journal of Political Economy* 105.6 (1997), pp. 1167–1200.

[23] Gordon L Clark, Andreas Feiner, and Michael Viehs. "From the stockholder to the stakeholder: How sustainability can drive financial outperformance". In: *Available at SSRN 2508281* (2015).

[24] Jose M Cruz. "Mitigating global supply chain risks through corporate social responsibility". In: *International Journal of Production Research* 51.13 (2013), pp. 3995–4010.

[25] Manuela Dantas. "Are ESG Funds More Transparent?" In: *Available at SSRN 3269939* (2021).

[26] Diane Del Guercio and Paula A Tkac. "Star power: The effect of monrningstar ratings on mutual fund flow". In: *Journal of Financial and Quantitative Analysis* 43.4 (2008), pp. 907–936.

[27] Jeroen Derwall, Nadja Guenster, Rob Bauer, and Kees Koedijk. "The eco-efficiency premium puzzle". In: *Financial Analysts Journal* 61.2 (2005), pp. 51–63.

[28] Dolf Diemont, Kyle Moore, and Aloy Soppe. "The downside of being responsible: Corporate social responsibility and tail risk". In: *Journal of Business Ethics* 137.2 (2016), pp. 213–229.

[29] Steven Dolvin, Jon Fulkerson, and Anna Krukover. "Do "good guys" finish last? The relationship between Morningstar sustainability ratings and mutual fund performance". In: *The Journal of Investing* 28.2 (2019), pp. 77–91.

[30] Himangshu Dutta. “Insights into the impacts of four current environmental problems on flying birds”. In: *Energy, Ecology and Environment* 2.5 (2017), pp. 329–349.

[31] Martin Eling and Roger Faust. “The performance of hedge funds and mutual funds in emerging markets”. In: *Journal of Banking & Finance* 34.8 (2010), pp. 1993–2009.

[32] Edwin J Elton and Martin J Gruber. “Mutual funds”. In: *Handbook of the Economics of Finance*. Vol. 2. Elsevier, 2013, pp. 1011–1061.

[33] Otto Eurola. “Key performance drivers of stocks in varying market states”. In: (2022).

[34] Mutual Fund Factbook. *Investment company institute*. 2002.

[35] Eugene F Fama and Kenneth R French. “Common risk factors in the returns on stocks and bonds”. In: *Journal of financial economics* 33.1 (1993), pp. 3–56.

[36] Eugene F Fama and James D MacBeth. “Risk, return, and equilibrium: Empirical tests”. In: *Journal of political economy* 81.3 (1973), pp. 607–636.

[37] Miguel A Ferreira, Aneel Keswani, Antonio F Miguel, and Sofia B Ramos. “The flow-performance relationship around the world”. In: *Journal of Banking & Finance* 36.6 (2012), pp. 1759–1780.

[38] Ryan Flugum and Matthew E Souther. “Stakeholder Value: A Convenient Excuse for Underperforming Managers?” In: *Available at SSRN 3725828* (2021).

[39] Milton Friedman. “A Friedman doctrine: The social responsibility of business is to increase its profits”. In: *The New York Times Magazine, 13* (1970), pp. 32–33.

[40] Paul Gompers, Joy Ishii, and Andrew Metrick. “Corporate governance and equity prices”. In: *The quarterly journal of economics* 118.1 (2003), pp. 107–156.

[41] Alexei P Goriaev, Theodoor Evert Nijman, Bas JM Werker, et al. *The dynamics of the impact of past performance on mutual fund flows*. Citeseer, 2002.

[42] Mark Grinblatt and Sheridan Titman. “The persistence of mutual fund performance”. In: *The Journal of finance* 47.5 (1992), pp. 1977–1984.

[43] Martin J Gruber. “Another puzzle: The growth in actively managed mutual funds”. In: *The journal of finance* 51.3 (1996), pp. 783–810.

[44] Russell L Gruen, Julian H Elliott, Monica L Nolan, Paul D Lawton, Anne Parkhill, Cameron J McLaren, and John N Lavis. “Sustainability science: an integrated approach for health-programme planning”. In: *The Lancet* 372.9649 (2008), pp. 1579–1589.

[45] Samuel M Hartzmark and Abigail B Sussman. “Do investors value sustainability? A natural experiment examining ranking and fund flows”. In: *The Journal of Finance* 74.6 (2019), pp. 2789–2837.

[46] Darryl Hedricks, Jayendu Patel, and Richard Zeckhauser. “Hot Hands in Mutual Funds: Short-Run Persistence of Performance in Relative Performance, 1974-1988”. In: *The Journal of Finance (March 1993)* (1993), pp. 93–130.

[47] Witold Henisz, Tim Koller, and Robin Nuttall. “Five ways that ESG creates value”. In: (2019).

[48] Edgar G Hertwich and Glen P Peters. “Carbon footprint of nations: a global, trade-linked analysis”. In: *Environmental science & technology* 43.16 (2009), pp. 6414–6420.

[49] Chastity Heyward. “The growing importance of social responsibility in business”. In: *Forbes, available at: https://www. forbes. com/sites/forbesbusinesscouncil/2020/11/18/the-growingimportance-ofsocial-responsibility-in-business* (2020).

[50] Harrison Hong and Marcin Kacperczyk. “The price of sin: The effects of social norms on markets”. In: *Journal of financial economics* 93.1 (2009), pp. 15–36.

[51] Jennifer Huang, Kelsey D Wei, and Hong Yan. “Participation costs and the sensitivity of fund flows to past performance”. In: *The journal of finance* 62.3 (2007), pp. 1273–1311.

[52] Israelsen. “A refinement to the Sharpe ratio and information ratio”. In: *Journal of Asset Management* 5.6 (2005), pp. 423–427.

[53] Israelsen. “Sharpening the Sharpe ratio”. In: *Financial Planning* 33.1 (2003), pp. 49–51.

[54] Archana Jain, Pankaj K Jain, and Zabihollah Rezaee. “Value-relevance of corporate social responsibility: Evidence from short selling”. In: *Journal of Management Accounting Research* 28.2 (2016), pp. 29–52.

[55] Michael C Jensen. “The performance of mutual funds in the period 1945-1964”. In: *The Journal of finance* 23.2 (1968), pp. 389–416.

[56] J Dave Jobson and Bob M Korkie. “Performance hypothesis testing with the Sharpe and Treynor measures”. In: *Journal of Finance* (1981), pp. 889–908.

[57] P Justice and J Hale. “Morningstar Sustainability Rating”. In: *Morningstar, https://s21. q4cdn. com/198919461/files/do Sustainability-Rating-Methodology. pdf downloaded* 4 (2016), p. 2020.

[58] Leora Klapper, Vıctor Sulla, and Dimitri Vittas. “The development of mutual funds around the world”. In: *Emerging Markets Review* 5.1 (2004), pp. 1–38.

[59] Sakis Kotsantonis and George Serafeim. “Four things no one will tell you about ESG data”. In: *Journal of Applied Corporate Finance* 31.2 (2019), pp. 50–58.

[60] Hao Liang and Luc Renneboog. “Corporate social responsibility and sustainable finance: A review of the literature”. In: *European Corporate Governance Institute–Finance Working Paper* 701 (2020).

[61] Hans Lööf, Andreas Stephan, et al. *The Impact of ESG on Stocks’ Downside Risk and Risk Adjusted Return.* Myndigheten för tillväxtpolitiska utvärderingar och analyser, 2019.

[62] RBC Wealth Management. “Women are leading the charge for Environmental, Social and Governance (ESG) investing in the U.S. amid growing demand for responsible investing solutions”. In: (2021). URL: https://www.rbcwealthmanagement.com/en-us/newsroom/2021-04-06/women-are-leading-the-charge-for-environmental-social-and-governance-esg-investing-in-the-us-amid-growing-demand-for-responsible-investing-solutions.

[63] Harry M Markowitz. “Portfolio Selection, 1959”. In: *Journal of Finance* 7 (1952), p. 7791.

[64] Bert Metz, Mike Hulme, et al. *Climate policy options post-2012: European strategy, technology and adaptation after Kyoto*. Routledge, 2013.

[65] Julia E Moore, Alekhya Mascarenhas, Julie Bain, and Sharon E Straus. “Developing a comprehensive definition of sustainability”. In: *Implementation Science* 12.1 (2017), pp. 1–8.

[66] Giacomo Morelli and Rita D’Ecclesia. “Responsible investments reduce market risks”. In: *Decisions in Economics and Finance* 44.2 (2021), pp. 1211–1233.

[67] Milton Moskowitz. “Choosing socially responsible stocks”. In: *Business and society review* 1.1 (1972), pp. 71–75.

[68] Sheila M Olmstead and Robert N Stavins. “Three key elements of a post-2012 international climate policy architecture”. In: *Review of Environmental Economics and Policy* (2012).

[69] Roger Otten and Dennis Bams. “European mutual fund performance”. In: *European financial management* 8.1 (2002), pp. 75–101.

[70] Harmony J Palmer. “Corporate social responsibility and financial performance: does it pay to be good?” In: (2012).

[71] Jacques Pézier, Anthony White, et al. *The relative merits of investable hedge fund indices and of funds of hedge funds in optimal passive portfolios*. Tech. rep. Henley Business School, University of Reading, 2006.

[72] Anna Prior. “Investors follow their conscience.” In: *The Wall Street Journal* (2015). URL: https://www.wsj.com/articles/investors-follow-their-conscience-1437154588.

[73] Enola Proctor, Douglas Luke, Annaliese Calhoun, Curtis McMillen, Ross Brownson, Stacey McCrary, and Margaret Padek. “Sustainability of evidence-based healthcare: research agenda, methodological advances, and infrastructure support”. In: *Implementation Science* 10.1 (2015), pp. 1–13.

[74] Arno Riedl and Paul Smeets. “Social preferences and portfolio choice”. In: (2017).

[75] Darrell Rigby and Suzanne Tager. “Learning the advantages of sustainable growth”. In: *Strategy & Leadership* (2008).

[76] TRV Risk. “Fund performance during market stress–The Corona experience”. In: (2022).

[77] Martin Rohleder. “The relation between past flows and future performance: simple investment strategies in the mutual fund sector”. In: *International Journal of Financial Studies* 3.1 (2015), pp. 3–30.

[78] K Geert Rouwenhorst. “The origins of mutual funds”. In: *Available at SSRN 636146* (2004).

[79] Vanya Rusinova and Georg Wernicke. “Short Selling and Performance on Corporate Social Responsibility: Evidence from a Natural Experiment”. In: *Academy of Management Proceedings*. Vol. 2019. 1. Academy of Management Briarcliff Manor, NY 10510. 2019, p. 17145.

[80] DODA Sanie and Albania Durres. “The importance of corporate social responsibility”. In: *Journal of Sociological Research* 6.1 (2015), pp. 86–91.

[81] Benjamin Schiller. “Never Mind Corporate Responsibility, Companies Can Solve Actual Social Problems.” In: *Fast Company* (2015).

[82] Robert C Scott and Philip A Horvath. “On the direction of preference for moments of higher order than the variance”. In: *The Journal of finance* 35.4 (1980), pp. 915–919.

[83] William F Sharpe. “Morningstar’s risk-adjusted ratings”. In: *Financial Analysts Journal* 54.4 (1998), pp. 21–33.

[84] William F Sharpe. “The sharpe ratio”. In: *Journal of portfolio management* 21.1 (1994), pp. 49–58.

[85] SIFMA. “Who Owns Stocks in America.” In: *SIFMA Insights* (2021).

[86] Erik R Sirri and Peter Tufano. “Costly search and mutual fund flows”. In: *The journal of finance* 53.5 (1998), pp. 1589–1622.

[87] Matthew Spiegel and Hong Zhang. “Mutual fund risk and market share-adjusted fund flows”. In: *Journal of Financial Economics* 108.2 (2013), pp. 506–528.

[88] Meir Statman and Denys Glushkov. “The wages of social responsibility”. In: *Financial Analysts Journal* 65.4 (2009), pp. 33–46.

[89] Linda Steg and Charles Vlek. “Encouraging pro-environmental behaviour: An integrative review and research agenda”. In: *Journal of environmental psychology* 29.3 (2009), pp. 309–317.

[90] Janet K Swim, Susan Clayton, and George S Howard. “Human behavioral contributions to climate change: psychological and contextual drivers.” In: *American Psychologist* 66.4 (2011), p. 251.

[91] Margarita Tsoutsoura. “Corporate social responsibility and financial performance”. In: (2004).

[92] UNCTAD. “The rise of the sustainable fund market and its role in financing sustainable development”. In: *United Nations* (2021).

[93] Gregory Unruh, David Kiron, Nina Kruschwitz, Martin Reeves, Holger Rubel, and Alexander Meyer Zum Felde. “Investing for a sustainable future: Investors care more about sustainability than many executives believe”. In: *MIT Sloan Management Review* 57.4 (2016).

[94] Wayne Visser. *The age of responsibility: CSR 2.0 and the new DNA of business.* John Wiley & Sons, 2011.

[95] Vincent A Warther. "Aggregate mutual fund flows and security returns". In: *Journal of financial economics* 39.2-3 (1995), pp. 209–235.

[96] Russ Wermers. *Momentum investment strategies of mutual funds, performance persistence, and survivorship bias.* Tech. rep. working paper, University of Colorado, 1997.

[97] Goetzmann William and RG Ibbotoson. "Do winners repeat? Patterns in mutual fund performance". In: *Journal of Portfolio Management* 20.9 (1994), p. 2.

[98] Yuchao Xiao, Robert Faff, Philip Gharghori, and Darren Lee. "An empirical study of the world price of sustainability". In: *Journal of business ethics* 114.2 (2013), pp. 297–310.

A Appendix

Table 1: This table reports the descriptive statistics of the 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating (⋆), the 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating (⋆⋆⋆⋆⋆), and the total 724 US open-end mutual funds, the Fama and French (1993) 3-factors, and the momentum factor over the period January 2000 - February 2021.

	Rating	Mean	St. dev.	Skew.	Kurt.	Min.	Median	Max.
Mutual funds		0.706	4.310	-0.664	1.807	-18.109	1.163	12.129
Low ESG funds	⋆	0.709	4.436	-0.672	2.298	-18.134	1.162	14.317
High ESG funds	⋆⋆⋆⋆⋆	0.703	4.251	-0.651	1.538	-18.088	1.277	11.035
Market proxy		0.555	4.555	-0.515	1.039	-17.230	1.135	13.650
SMB^*		0.286	3.266	0.753	9.690	-17.290	0.205	21.480
HML^*		0.113	3.283	0.095	2.958	-13.920	-0.085	12.480
MOM^*		2.839	2.435	0.920	-0.027	0.000	2.200	9.600

Note: All indices are analyzed on basis of excess returns, unless indicated with an asterisk (*).

Table 2: This table reports adjusted R^2 for equally weighted portfolio, estimated with (1) the Capital Asset Pricing Model, (2) the Fama and French (1993), and (3) the Carhart (1997) model over the period January 2000 - February 2021, and the minimum and maximum Adjusted R^2, the 25%, and 75% quantile and the median R^2 for the individual funds. Panel A presents the results for the 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating, Panel B reports results for the 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and Panel C reports results for the total 724 US open-end mutual funds.

	Equally weighted portfolio	Individual funds				
		Min.	25% Quantil	Median	75% Quantil	Max.
Panel A: Low ESG funds						
CAPM	0.941	0.177	0.674	0.743	0.848	0.974
Fama and French	0.968	0.178	0.735	0.849	0.916	0.975
Carhart	0.968	0.180	0.735	0.849	0.916	0.849
Panel B: High ESG funds						
CAPM	0.940	-0.005	0.663	0.767	0.822	0.972
Fama and French	0.953	0.004	0.708	0.818	0.890	0.818
Carhart	0.953	0.008	0.708	0.819	0.894	0.819
Panel C: All funds						
CAPM	0.952	-0.002	0.672	0.751	0.838	0.975
Fama and French	0.964	0.008	0.729	0.838	0.901	0.986
Carhart	0.964	0.012	0.729	0.838	0.902	0.986

Table 3: This table reports alphas for equally weighted portfolio, estimated with (1) the Capital Asset Pricing Model, (2) the Fama and French (1993), and (3) the Carhart (1997) model over the period January 2000 - February 2021, the minimum and maximum alpha, the 25%, and 75% quantile, the median alpha, and the α distribution for the individual funds. Panel A presents the results for the 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating, Panel B reports results for the 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and Panel C reports results for the total 724 US open-end mutual funds.

	Equally weighted portfolio		Individual funds					Alpha distribution		Test
	Alpha	*t*-stat	Min.	25% Quantil	Median	75% Quantil	Max.	sign. < 0 (%)	sign.> 0 (%)	Z-test
Panel A: Low ESG funds										
CAPM	0.054	0.794	-0.793	-0.051	0.071	0.175	0.620	5.120	17.169	1.360
Fama and French	0.004	0.075	-0.779	-0.115	0.018	0.133	0.468	6.928	13.554	2.514**
Carhart	-0.003	-0.042	-1.172	-0.090	0.001	0.137	0.675	9.036	9.036	0.753
Panel B: High ESG funds										
CAPM	0.070	1.069	-0.474	-0.114	0.717	0.223	0.778	4.337	21.173	
Fama and French	0.051	0.871	-0.459	-0.116	0.061	0.193	0.715	7.653	20.663	
Carhart	-0.007	-0.078	-0.714	-0.188	-0.009	0.148	1.310	11.224	10.714	
Panel C: All funds										
CAPM	0.063	1.047	-0.793	-0.096	0.071	0.199	0.778	5.110	19.475	
Fama and French	0.029	0.566	-0.779	-0.115	0.038	0.165	0.715	7.735	17.265	
Carhart	-0.005	0.066	-1.172	-0.157	-0.004	0.142	0.131	11.740	8.702	

Note: Statistical significance at 1, 5, and 10% levels are indicated by ***, **, and * respectively.

Table 4: This table reports alphas for equally weighted portfolio, estimated with (1) the Capital Asset Pricing Model, (2) the Fama and French (1993), and (3) the Carhart (1997) model over the sub periods April 2000 - August 2002, September 2002 - October 2007, November 2007 - February 2009, March 2009 - January 2020, and February 2020 - February 2021, the minimum and maximum alphas, the 25%, and 75% quantile, the median alphas, and the α distribution for the individual funds. Panel A presents the results for the 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating, Panel B reports results for the 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and Panel C reports results for the total 724 US open-end mutual funds.

	Equally weighted portfolio		Individual funds					Alpha distribution	
	Alpha	t-stat	Min.	25% Quantil	Median	75% Quantil	Max.	sign. < 0 (%)	sign.> 0 (%)
Subperiod: April 2000 - August 2002									
Panel A: Low ESG funds									
CAPM	0.556**	2.137	-2.228	-0.053	0.416	1.138	3.376	6.325	20.181
Fama and French	-0.020	0.112	-2.298	-0.440	0.032	0.348	2.624	14.157	15.361
Carhart	-0.330	0.997	-5.724	-0.749	-0.362	0.206	5.127	18.976	4.819
Panel B: High ESG funds									
CAPM	0.431*	1.975	-1.576	-0.146	0.296	0.915	3.373	3.571	18.112
Fama and French	0.076	0.403	-1.930	-0.491	0.104	0.534	3.272	9.949	10.969
Carhart	-0.454	-1.366	-2.962	-1.000	-0.368	0.203	2.507	16.837	2.806
Panel C: All funds									
CAPM	0.493**	2.152	-2.228	-0.078	0.356	1.059	3.376	3.591	19.200
Fama and French	0.032	0.179	-2.298	-0.453	0.076	0.436	3.272	12.155	10.912
Carhart	-0.397	-1.237	-5.724	-0.905	-0.362	0.205	5.127	17.541	4.834
Subperiod: September 2002 - October 2007									
Panel A: Low ESG funds									
CAPM	0.281*	2.673	-0.446	0.015	0.191	0.459	2.092	2.108	27.711
Fama and French	0.200**	2.579	-0.483	-0.060	0.101	0.383	1.865	2.711	21.988
Carhart	0.136	1.258	-0.722	-0.116	0.047	0.223	1.626	6.024	18.675
Panel B: High ESG funds									
CAPM	0.322***	2.706	-0.483	-0.003	0.153	0.526	2.188	0.000	17.857
Fama and French	0.300***	2.924	-0.707	-0.018	0.153	0.445	2.187	0.510	16.837
Carhart	0.256*	1.779	-0.952	-0.065	0.051	0.378	2.792	1.531	15.561
Panel C: All funds									
CAPM	0.303***	2.830	-0.483	0.005	0.176	0.490	2.188	0.552	13.950
Fama and French	0.254***	2.889	-0.707	-0.037	0.138	0.404	2.187	0.414	14.917
Carhart	0.201	1.632	-0.952	-0.092	0.049	0.337	2.792	1.796	11.464
Subperiod: November 2007 - February 2009									
Panel A: Low ESG funds									
CAPM	0.043	0.140	-2.198	-0.308	-0.028	0.526	1.660	10.241	11.446
Fama and French	-0.075	0.234	-3.070	-0.370	-0.049	0.309	1.515	13.253	8.133
Carhart	-0.218	-0.439	-2.887	-0.764	-0.272	0.365	4.487	22.590	5.422
Panel B: High ESG funds									
CAPM	0.357	0.667	-2.965	-0.102	0.316	0.737	2.492	4.847	18.622
Fama and French	0.222	0.547	-3.050	-0.056	0.242	0.569	1.907	4.592	19.133
Carhart	0.169	0.267	-4.241	-0.512	0.282	0.995	4.888	7.653	12.755
Panel C: All funds									
CAPM	0.213	0.511	-2.965	-0.233	0.165	0.618	2.492	5.663	16.989
Fama and French	0.086	0.242	-3.070	-0.253	0.111	0.436	1.907	6.768	14.503
Carhart	-0.008	-0.015	-4.241	-0.631	0.039	0.706	4.888	8.011	10.773

Subperiod: March 2009 - January 2020									
Panel A: Low ESG funds									
CAPM	-0.243^{***}	-3.468	-2.035	-0.358	-0.228	-0.059	0.475	30.422	0.000
Fama and French	-0.168^{***}	-2.841	-1.710	-0.283	-0.118	0.025	0.445	25.301	1.807
Carhart	-0.172^{*}	-1.927	-2.307	-0.314	-0.147	0.033	0.634	24.398	4.217
Panel B: High ESG funds									
CAPM	-0.129^{*}	-1.724	-0.835	-0.284	-0.107	0.046	0.474	25.000	1.020
Fama and French	-0.144^{**}	-2.048	-0.924	-0.310	-0.095	0.051	0.438	26.531	1.020
Carhart	-0.203^{*}	-1.908	-1.533	-0.350	-0.142	0.008	0.548	29.592	1.785
Panel C: All funds									
CAPM	-0.181^{***}	-2.795	-2.035	-0.323	-0.152	-0.015	0.475	25.414	2.901
Fama and French	-0.155^{**}	-2.502	-1.710	-0.289	-0.103	0.043	0.445	23.900	1.519
Carhart	-0.189^{**}	-2.016	-2.307	-0.329	-0.146	0.018	0.634	22.790	1.657
Subperiod: February 2020 - February 2021									
Panel A: Low ESG funds									
CAPM	-0.056	-0.086	-3.214	-0.811	-0.375	0.630	5.582	9.639	19.277
Fama and French	0.332	0.728	-2.946	-0.360	0.175	0.865	6.523	7.530	22.289
Carhart	1.020	1.232	-4.125	-0.193	0.902	1.705	9.364	5.723	21.988
Panel B: High ESG funds									
CAPM	0.135	0.412	-3.119	-0.403	0.099	0.565	3.183	5.357	12.500
Fama and French	0.085	0.170	-4.42-	-0.349	0.072	0.504	3.154	8.673	11.990
Carhart	0.792	0.865	-4.714	0.066	0.662	1.542	6.412	5.612	17.092
Panel C: All funds									
CAPM	0.047	0.110	-3.214	-0.568	-0.087	0.594	5.582	7.044	12.293
Fama and French	0.198	0.456	-4.420	-0.351	0.120	0.581	6.523	6.077	10.083
Carhart	0.900	1.147	-4.714	0.040	0.748	1.654	9.364	3.453	18.923

Note: Statistical significance at 1, 5, and 10% levels are indicated by ***, **, and * respectively.

Table 5: This table reports performance measurement values and their rankings for the 332 US open-end mutual funds with an one star Morningstar Sustainalytics rating, the 392 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and all the 724 US open-end mutual funds, the market proxy, the Small minus Big (SMB) factor, the High minus Low (HML) factor, and the Momentum (MOM) factor over the period January 2000 - February 2021.

	Panel A: Measurement Value					Panel B: Ranking					Panel C: Test
	Sharpe ratio	Modified Sharpe ratio (Israelsen)	Adjusted Sharpe ratio (Pezier, White et al. (2006))	Sortino ratio	Treynor ratio	Sharpe ratio	Modified Sharpe ratio (Israelsen)	Adjusted Sharpe ratio (Pezier, White et al. (2006))	Sortino ratio	Treynor ratio	Jobson & Korkie (1981) Test
All funds	0.133	0.133	0.132	0.179	-0.231	3	3	3	3	6	0.143
Low funds	0.130	0.130	0.128	6.957	-0.223	4	4	4	1	5	0.335
High funds	0.135	0.135	0.133	0.170	-2.411	2	2	2	4	7	-
Market proxy	0.122	0.122	0.121	0.161	-0.110	5	5	5	5	3	2.404***
SMB	0.088	0.088	0.088	0.132	-0.191	6	6	6	6	4	1.106
HML	0.034	0.034	0.034	0.050	0.032	7	7	7	7	2	1.445*
MOM	1.166	1.166	1.574	1.976	1.347	1	1	1	2	1	-9.531***

Note: Statistical significance at 1, 5, and 10% levels are indicated by ***, **, and * respectively.

Table 6: This table reports means and corresponding standard deviations of the absolute flows ($ mln), relative flows (%), performance (Jensen's α), size ($ mln) and age (months) for the 309 US open-end mutual funds with an one star Morningstar Sustainalytics rating, the 359 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and all the 668 US open-end mutual funds over the period March 2018 - February 2021.

	Low rated funds		High rated funds		All funds	
	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
Absolute flow, $ mln	-9.200	60.799	-2.750	70.786	-5.732	66.431
Relative flow, %	-0.606	28.273	0.662	35.617	1.095	89.738
Performance	-0.551	1.645	0.083	1.085	-0.210	1.409
Total Net Assets, $ mln	1,160.489	6,136.226	921.489	2,311.369	1,032.044	4,505.864
Age, months	23.936	11.214	213.914	120.546	126.035	129.769

Table 7: This table reports coefficient estimates and the corresponding standard errors of the two stage Fama and MacBeth (1973) regression of Equation 18 for the 309 US open-end mutual funds with an one star Morningstar Sustainalytics rating, the 359 US open-end mutual funds with a five star Morningstar Sustainalytics rating, and all the 668 US open-end mutual funds over the period March 2018 - February 2021. The dependent variable is fund net relative flow. The explanatory variables include a constant, the lagged relative flow, a dummy variable, the log of fund size, the log of fund age, and Jensen's alpha, estimated over a period of six months.

	Coef.	S.e.
γ_0	0.003	0.062
$f_{i,t-1}$	-0.052	0.069
$D_{i,t}$	-0.020	0.067
$D_{i,t}$ x $f_{i,t-1}$	-0.001	0.231
$\alpha_{i,t}$	0.396	1.913
$D_{i,t}$ x $\alpha_{i,t}$	-0.558	2.210
$log(TNA_{i,t-1})$	-0.423	0.647
$log(Age_{i,t-1})$	-0.030	0.105

Note: Statistical significance at 1, 5, and 10% levels are indicated by ***, **, and * respectively.