Norwegian School of Economics
Bergen, Fall 2022

Analysing Fund Flow in the European Sustainable Mutual Fund Market

An Empirical Analysis of Sustainability-rated European Listed Mutual Funds before and after the Covid-19 recession

Mats Hellan & Sander Sørensen

Supervisor: Andreas Ørpetveit

Master thesis, Economics and Business Administration

Major: Finance

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis has been written as part of our Master of Science degree in Economics and Business Administration at Norwegian School of Economics (NHH). The process of writing has been both challenging and memorable. Mostly memorable, as it has been very interesting to learn about a topic of our own choice.

We would like to extend our gratitude towards our supervisor, Andreas Ørpetveit, for his counselling and guidance. His knowledge and ideas have helped us immensely along the way.

Abstract

We analyse the effect of the Morningstar Sustainability Rating on fund flow in the European mutual fund market. The Covid-19 recession in March 2020 was a pivot for the industry, and we look at one-year periods before and after. Through OLS and fixed effects methods, we find that investors value the Morningstar Sustainability Rating, as higher ratings are associated with higher fund flow. The results show an increased impact of sustainability ratings following the recession, suggesting a change in fund flow dynamics. Furthermore, through a difference-in-difference model we provide evidence that flow to High-rated funds increased following the recession and that the High-rated funds have relatively higher inflow compared to Low-rated funds. The impact on each domicile varies. More than half of the sample countries have higher expected fund flow for High-rated funds than Low-rated funds throughout the entire period. However, only a small portion of the countries have increased fund flow to High-rated funds following the Covid-19 recession.

Keywords: *Mutual funds, Fund flow, Sustainability, ESG, Morningstar Sustainability Rating, Covid-19*

Contents

List of Figures

List of Tables

1 Introduction

Mutual funds have become increasingly popular in recent decades, as they offer investors the expertise of professional fund managers, combined with diversification and liquidity (SEC, 2022). The prevalence continues in today's financial markets, as total net assets in European mutual funds have more than doubled since 2015 (ICI, 2022). Similarly, Norwegian fund investments peaked in May 2021, where 46 percent of the population owned shares in mutual funds, excluding pension arrangements (VFF, 2021). Contextualised, that is a 10 percent increase from 2018, suggesting more people are aware of and are actively investing in mutual funds.

Sustainability has recently received increased amounts of attention in the mutual fund market. One of the most popular approaches to sustainable investing has been third-party ESG scores created by independent analytics companies such as MSCI, Morningstar, Refinitiv, Bloomberg, and FTSE (MSF, 2022). Especially the Morningstar Sustainability Rating has seen increased popularity since its introduction in 2016, now becoming one of the essential tools for responsible mutual fund investing. It serves as a crucial metric when determining the industry-relative ESG risk, while also being an indicator of overall sustainability.

According to (Bialkowski & Starks, 2016; Hartzmark & Sussman, 2019), the shift towards green investments has gradually increased over the last decade and points to several reasons why sustainability is preferred. They argue that investors believe higher sustainability implies that risk-adjusted returns are better or that investors' financial motivation is less of a priority. Additionally, increased awareness among investors has an effect, which is likely to continue growing as more challenges regarding climate change and other ESG aspects emerge (PRI, 2021).

Several studies have analysed the connection between sustainability and funds. However, there is an apparent shortage in recent literature and research touching upon the connection between fund flow and sustainability. In that regard, we wish to shed light on how investors allocate their funds. We found this idea quite intriguing as Pástor & Vorsatz (2020) suggest that High-rated funds received relatively higher fund flows in the post-covid recovery period and therefore points to sustainability being considered a necessity, not a luxury, by investors. Additionally, Hartzmark & Sussman (2019) find that mutual fund investors collectively regard sustainability as essential when allocating capital to funds.

The Covid-19 outbreak caused a sudden, sharp collapse in global financial markets known as the 2020 Market Crash. Following the Covid-19 crisis, several assertions analysed the relationship between sustainability and fund flow. Many concluded that ESG elements provide investors with a safeguard from downside risk in challenging market conditions. However, there is currently little research on how the fund flow of sustainability-rated funds has changed over time. From the steady bull run leading up to the Covid-19 crisis to after the market started to re-balance and recover from the recession in March 2020. This development is interesting as it could provide valuable insight into investor preferences regarding sustainable products.

Europe as a market is well-developed within ESG integration, and we have chosen to focus on European countries with well-established mutual fund markets in this thesis. Our motivation for this scope is primarily to look at Europe and uncover trends within sustainable investing in the Eurozone and European financial markets, which differs from most studies in the past that research the US market. With that in mind, our first research question is as follows:

(1) Do investors value the Morningstar Sustainability Rating?

When considering the results from previous studies conducted by (Hartzmark & Sussman, 2019), investors seem to value sustainability. Based on these results, and the increasing worldwide interest in ESG-related topics (Google, 2022), our hypothesis for research question (1) is that investors value Morningstar Sustainability Rating greatly when investing. We expect the results to show more inflow to higher-rated funds and that the net flow increases exponentially with higher ratings.

Furthermore, we are interested in capturing any development or change in the flow-to-sustainability relationship that could increase our understanding of the investment climate throughout Europe. Our research question to be answered is:

(2) How does Morningstar Sustainability Rating affect fund flow before and during the recovery period following the Covid-19 market crisis?

In the period following the covid recession, sustainability gained excessive media attention.

Additionally, there was a flight-to-safety effect among investors, causing further market disruptions. Since higher-rated ESG assets have outperformed the general market over time (ref. 3.2.2), we believe that investors saw the Covid-19 recession as an opportunity to increase their green market exposure. In line with our hypothesis for research question (1), we believe investors value sustainability. Hence, our predicted outcome and hypothesis for research question (2) is a shift towards more sustainable investments in the post-covid recovery period.

The sample contains data from European domiciles, whereas institutions and investors have different demographic traits. To further explore potential differences within these domiciles, which in turn may contribute to biased results when analysing all regions collectively, we want to look at the Morningstar Sustainability Rating's relation to fund flow in each region separately. This is considered a sub-question of research question (2), and will follow a more explorative approach than the others. The goal is to analyse regional trends of sustainable investing and uncover differences within European countries. The research question to guide this is:

> *(3) Are there differences in fund flow and the flow-to-sustainability relationship within the different domiciles?*

Europe as a continent is vast and has regions of both wealth and relative poverty. Therefore, we expect behavioural differences in institutional and retail investors in different regions. Considering this, we have developed the following hypothesis. In general, we believe European countries with stable and developed economies, that score low (good) on the Morningstar Sustainability Atlas (Section 3.4), will have significantly greater fund flow to higher-rated funds. In contrast, the difference between higher- and lower-rated funds is expected to be smaller in less developed countries and countries with higher scores.

The results from using both OLS and fixed effects models, testing rating as a discrete variable and looking at the rating factorised, confirm that a better rating is associated with a higher flow of funds to the respective mutual fund. Hence, confirming our hypothesis that investors value the Morningstar Sustainability Rating.

An analysis using difference-in-difference methodology was conducted to answer the second research question. The results suggested that the effect of the Morningstar Sustainability

Rating increased during the recovery period. High-rated funds had a treatment effect, or increased expected fund flow, of 1.00 percent. Combined with the OLS and Fixed effects results, fund flow seems to be more affected by the Morningstar Sustainability Rating after the Covid-19 crisis, and High-rated funds are the ones that benefit the most from this increased effect.

Furthermore, we believed the effect of sustainability ratings within and between countries would show significant differences when considering the general demographic and socioeconomic differences between European domiciles. The results showed that several countries had increased net flow to High-rated funds compared to Low-rated funds. However, only difference-in-difference results from Great Brittan, France, Italy, and Lichtenstein were statistically significant at the 0.05 level, where the latter even had a negative coefficient. Suggesting the results were less convincing than our hypothesis predicted. It is difficult to identify any clear pattern of the flow to sustainable funds within each domicile between the two periods.

The structure for this thesis is as follows: In Section 2, we will review the current literature on mutual fund flow and sustainability. Section 3 describes the background and our motivation for writing about the topic. Next is Section 4, a chapter that presents our data, variables, and summary statistics for the dataset. Followed by our methodology in Section 5. Empirical results are presented in Section 6, and the discussion is in Section 7. Finally, we present our conclusion in Section 8, with references and the Appendix in the following sections.

2 Litterature Review

This chapter will give insight into previous studies on sustainability and fund flow. Many of the highlighted reports address the impact of Morningstar Sustainability Rating on mutual funds and to what degree investors value sustainability within several periods. These articles are essential for developing our approach towards the methodology and applying new perspectives to existing research by analysing a newer timeframe and other regions. Our goal is to contribute additional perspectives to this field of research.

2.1 Mutual Fund Flow

Sirri & Tofano (1998) present a comprehensive analysis researching inflow and outflow within equity mutual funds. They provide evidence that investors flock to high-performing funds disproportionately while failing to flee mutual funds that are lower-performing. They also state that fund flows are sensitive towards fees, but the response from consumers is asymmetric as they act differently to higher or lower fee levels. Furthermore, the findings suggest that consumers respond to the degree of risk in their portfolios, which may disturb the managers´ incentives in terms of increasing fund volatility. Lastly, they studied how media coverage affects mutual fund growth. They found some evidence that a more significant proportion of media attention is related to faster growth within mutual funds.

The article examines mutual fund flow in general. Still, it provides essential perspectives that are key to understanding different dynamics within mutual fund investment that can relate to our findings within the sustainable mutual fund market.

2.2 Introduction of Morningstar Sustainability Rating

A few studies have examined the effect of Morningstar´s Sustainability Rating on mutual fund flow following its introduction to the market in March of 2016. These studies are highly relevant to our research in terms of a similar approach. However, they are conducted in the U.S. mutual fund market.

Firstly, Hartzmark & Sussmann (2019) analysed how mutual fund flow reacted to the ratings' publication. According to their study, causal evidence that investors value sustainability is present. High-rated funds experienced roughly 0.30 percent greater inflows, whilst Low-rated had a negative fund flow of 0.44 percent monthly compared to Average-rated funds when only including ratings. On the other hand, above-average and below-average funds were not statistically significant. As the globe ratings may vary due to different variables associated with flow, they include several controls, such as lagged monthly returns. Even after controlling for all additional variables, they find similar results, suggesting that investors primarily have responded to higher sustainability ratings. Amman et al. (2018) also published an article exploring the relationship between mutual fund flow and sustainability utilising MSR as the main explanatory variable. They examined approximately 1000 funds per month with 49 percent representation of retail share classes. They found that average-rated funds receive respectively 0.23 percent higher and 0.29 percent lower than Low- and High-rated funds per month. In particular, they found strong evidence that retail investors moved money from Low- to High-rated funds and argued that it is much lower for institutional investors.

Both articles have studied investor behaviour following the introduction of the globe system within a similar time frame in the U.S. mutual fund market. The literature indicates that Morningstar Sustainability Rating has indeed had an effect on fund flow for mutual funds that received a rating when introduced.

2.3 Covid-19 market crash

Pástor & Vorsatz (2020) published an article focusing on fund performance and flows during the covid-19 crisis. They found that most active funds were underperforming passive benchmarks during the Covid-19 recession. In contrast, mutual funds with high sustainability ratings performed relatively better. This has led to investors favouring high-rated funds when reallocating capital, receiving relatively larger aggregated fund flows in the recovery period compared to earlier. They also suggest that investors view sustainability as a necessity rather than a luxury good.

Furthermore, Ferriani & Natoli (2020) analysed ESG risks during covid, exploiting ESG to estimate effects on fund flow investing. These results were based on an approach utilising Low-rated versus High-rated funds to examine the impact. They indicate no significant differences in the pre-crash phase (Jan. 2020 – Feb. 2020), in contrast to the recovery phase where cumulated fund flow appears to be of more importance as the Low-rated dummy is significantly negative whilst the High-rated dummy is significantly positive, identifying a flight-to safety effect towards High-rated funds. They, therefore, argue that investors highly prefer sustainability-rated mutual funds, with environmental concerns as the top priority.

These findings are essential to our research and have been a great inspiration when designing our analysis.

3 Background

In this chapter, relevant topics will be presented to help better understand why our research questions are interesting to explore.

3.1 ESG

Within finance, sustainability is often referred to as Environmental, Social, and Governance (ESG). These factors are essential when evaluating a portfolio across several sustainability aspects and help investors align their ESG criteria in terms of sustainable development when considering potential investments (Boffo & Patalano, 2020). Simplified, ESG investing can be understood as market participants aiming to achieve a common goal called "green investing". Meaning stakeholders consider the environmental, social and governance dimensions at their core and aim to improve portfolios or companies within these dimensions (De Spiegeleer et al., 2020). Firstly, the environmental part mainly focuses on climate change, including factors such as company contribution to reducing greenhouse gas emissions, green energy, and waste management. Secondly, the social dimension includes workplace safety & health, human rights, and labour standards. Lastly, governance covers principles defining responsibility, expectation and rights between stakeholders governing corporations (Robeco, n.d.).

ESG investing has become increasingly popular over the years (Investopedia, 2022) as fear of climate change has led investors to value the impact of their money. However, it originates back to the 1960s within social responsibility, where some investors excluded companies from their portfolios if they had business activities/involvement linked to, for example, the South African apartheid regime (MSCI, n.d.). Environmental concerns and global warming started to receive international attention as late as the 1980s. James E. Hansen published one of the first assessments on how human emissions had significantly affected the global climate (1988) and testified this to congress. Later, the Kyoto Protocol was implemented in 1997 to reduce greenhouse gases globally (UN, n.d.). Aligned with the increasingly higher focus on environmental issues, ESG investments experienced steady growth, further accelerating in 2013-14 when studies showed that if a corporation has good financial results, it is often associated with a good sustainability performance (Kell, 2018). In addition, The UN Principles for Responsible Investments' foundation is built on ESG factors and is working to promote incorporating these dimensions into investment decision-making. By the end of March 2021, there were over 3,400 investors that managed over 121 trillion USD worth of assets under management that follow these principles as signatories (UN PRI, 2021). Lastly, ESG investments into responsible investment funds increased by over 200 percent from 2019 to 2020 and 35 percent from 2020 to 2021 (Investment Association, 2020 & 2021).

ESG factors have significantly impacted investor preferences during the last decade. We believe exploring ESG aspects within mutual funds is an exciting approach to better understanding the implications of sustainability in the financial markets.

3.2 Covid-19

3.2.1 World Health Emergency

The World Health Organization (WHO, 2020) proclaimed the Covid-19 epidemic a global public health emergency on January 30th, 2020, after receiving reports of more than 7,000 cases worldwide. The number of cases then started to rise, and by March 11th, the WHO (2020) declared Covid-19 a pandemic. As a result, many regions of the world went into lockdown. By April 2020, approximately half of the world's population was under lockdown (Sandford, 2020).

3.2.2 Market Crash & Sustainability

Prior to WHO's declaration, financial markets were in a bull phase and had experienced steady growth since the end of the financial crisis in 2009 (RBC, 2022). However, this changed as the dramatic collapse in the world stock markets, known as the 2020 Market Crash, came in response to the worldwide economic activity staggering following the lockdown, causing a market panic. The financial markets entered a bear market, which occurs if the index drops at least 20 percent from the previous peak (Gonzalez et al., 2005). Figure 3.1 depicts the MSCI World Index, reaching its pre-covid peak at 2,434.5 on the 12th of February 2020, and its lowest at 1,596 on the 23rd of March, a staggering decline of 34.4 percent. According to the International Monetary Fund, the Great Lockdown is the most significant economic crisis since the Great Depression (Gopinath, 2020). Additionally, the market collapse also saw the fastest fall in stock markets in recorded financial history (Li, 2020).

Although the decline was dramatic, it only took a few months for the market almost to recover completely. Emergency funding in response to the pandemic to stimulate the economy, and partly re-opening societies contributed to the quick recovery, which led to the MSCI exceeding previous highs by August 2020. It was recorded as the fastest recovery ever (Jasinski, 2020), further increasing towards April 2021.

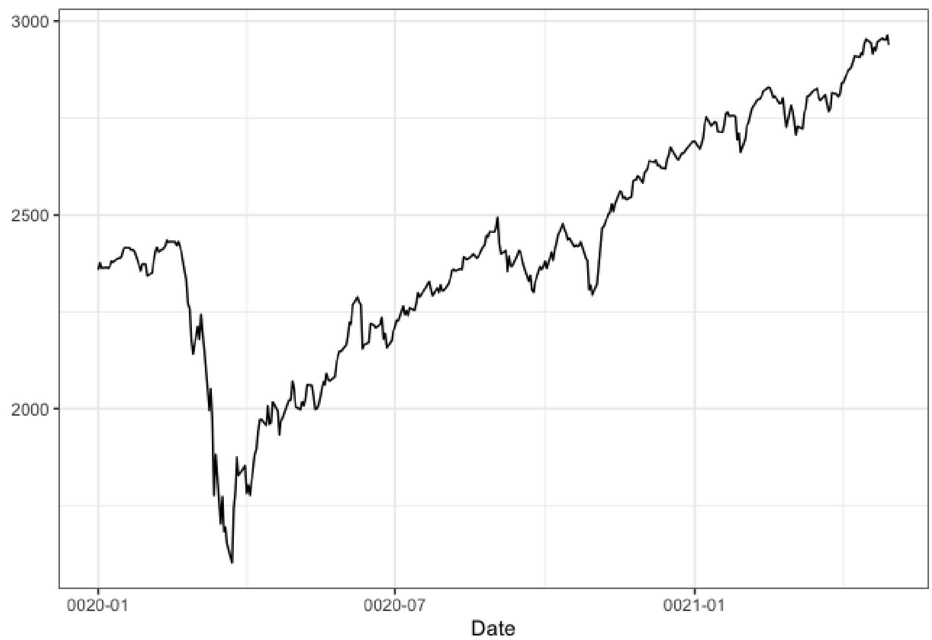


Figure 3.1: MSCI World Index
Shows the MSCI World Index from 01.01.2020 to 30.04.2021.

Furthermore, sustainability was important during the market crash, and several articles concerning ESG were published following the event. Some of these articles suggest that ESG exposure has worked as downside protection during the Covid-19 crisis, and that businesses with high ESG engagement have outperformed competitors with lower ESG exposure. The significance of ESG during the early covid-phase was emphasised as early as April 2020, where ESG behaved as an equity vaccine by outperforming other holdings (Morningstar, 2020). An article published in Financial Times confirms these findings and states that more than half of ethical and sustainable funds outperformed the MSCI index relative to steep downturns (Darbyshire, 2020). Another aspect of sustainable funds outperforming their peers is that portfolio companies are reaching towards something more than just profits, caring more for all their shareholders and serving a bigger purpose. Meaning they are potentially better equipped for a crisis as the long-term impact to create a sustainable world is the primary vision (Polman, 2020). These articles are examples of the increasing ESG interest seen early in the covid-phase. Through analysis, we explore whether this hype has converted into new fund flow dynamics following the recession.

3.3 The European Market

Our thesis is limited to selected European Countries. According to the Morningstar Sustainability Atlas (Morningstar, 2021a), the European continent, particularly northern countries, is the best region in terms of Portfolio Sustainability scores and ESG practices. Respectively, The Netherlands, France, and Sweden have the world´s most sustainable stock market, whereas Norway scores lowest within our sample. The figure shows the different scores based on the countries from our sample and some selected countries for comparison:

17.86-22.00	*22.01-24.03*	*24.04-27.42*	*27.43-30.00*	*30.01-42.95*
Denmark	*UK*	*Belgium*	*China*	*Russia*
Finland	*Austria*	*Norway*	*South Korea*	*Qatar*
Netherlands	*Switzerland*	*Japan*	*Turkey*	*Brazil*
Sweden	*Italy*	*New Zealand*	*Mexico*	
Ireland	*Australia*			
Spain	*Canada*			
Portugal	*US*			
Germany				
France				

Figure 3.2: Morningstar Sustainability Atlas
The grey header displays the Country Indexes Portfolio Sustainability Score (2021) in several intervals. The lower, the better. The countries within our samples are given a colour based on their Index range. Some countries from outside our sample are added to complement the index further.

While our sample has an average of 21.7, US and Canada have a combined score of 23, which could be an interesting market to include. However, we have decided to concentrate on the European market. In addition, the combination of ESG focus among the general European public and relatively equal Sustainability Score Indexes per country makes this market optimal for improving our understanding of the ESG factors. In contrast, regions such as Asia could also be interesting to explore. However, they experience lag and vary a lot more. This could lead to biases in our analysis as the basis of comparison differs and could potentially alter our results negatively.

The sample funds are distributed across several countries, but there is a noticeably higher concentration in Luxembourg than in all the others. A report from EY (2020) points to several reasons why more and more funds choose Luxembourg as their domicile, such as potential tax haven advantages, well-tested fund solutions, and developed regulatory frameworks that benefit the funds to be freely marketed throughout the EU.

4 Data & Summary Statistcs

The primary data source used in this thesis has been Morningstar Direct, an investment and portfolio analysis software issued by Morningstar. It contains an extensive database of mutual fund data, covering all the necessary variables. We collected raw data from 25,237 funds originating in 22 countries through Morningstar Direct Desktop. Within each fund, there is, if available, monthly data on fund size (Total Net Assets), Morningstar Sustainability Rating, performance (Return), and if it is an index fund or not. The main timeframe of the requested data is January 2015 to August 2022, which will be trimmed later on in the process of data cleaning to fit the desired scope.

The data output from the database is based on open-end mutual funds investing predominantly in Equity, including "dead funds" from the chosen domiciles within our timeframe. The following section will present a more detailed explanation of the data collection.

4.1 Data Collection Process

There were multiple steps involved in the process of extracting our desired data table. Firstly, we chose open-ended funds, which are mutual funds that can issue unlimited new shares and are priced daily based on their Net Asset Value / NAV (Investopedia, 2021a). Furthermore, the chosen domiciles were added in line with our scope for the European market. The global category group was set to equity, including mutual funds that invest in company shares. “Oldest Share Class” for open-end funds is the share class used in this analysis, meaning the total net assets are calculated from the share class with the most extended history within the mutual fund. However, there are cases where funds have launched several classes. To ensure that the dataset uses the appropriate share class for performance comparisons, Morningstar has developed a methodology that considers other share classes when appropriate (Morningstar, 2016).

Lastly, to maintain consistency, we have used Euro as the common currency in our data and included *dead funds*[1] to increase our data foundation and avoid survivorship bias (Investopedia, 2021b).

[1]Mutual funds that have gone bust. May result in an overestimated historical flow if excluded.

In the next part of extracting data, performance reporting with no grouping was used. Additional variables such as Fund Name, SecID, ISIN and Index Fund (Yes/No) were added. In addition, historical variables had to be included, consisting of Return, Fund Size comprehensive and the Morningstar Sustainability Rating. Following all these steps, a data table ready for extraction was produced. The processing of the collected data can be found in the appendix.

4.2 Sample period

The time frame of the original sample is between January 2015 and august 2022. The final data is trimmed to create adequate sample periods that correspond with the purpose of this research. Furthermore, we have used monthly observations for all the data, which in this research is natural due to Morningstar Sustainability Ratings only being published monthly.

The final sample period timeline is from February 2019 to April 2021, excluding March 2020, as we want to explore the changes before and after the event, not the event itself. We want to analyse the flow-rating relationship in a recent market environment where trends and patterns are closer to the financial climate at the time of writing this thesis. Furthermore, we want to examine how mutual fund flow has developed following a recession when market optimism is high, and the risk perspective has fewer nuances. This is especially interesting to research considering the ESG focus that persisted through the recovering market. To cover the immediate reactions, and the crux of the recovery following the market crash, we set the reference period to one year after the event, from April 2020 to April 2021.

To further understand the market changes for sustainability-rated funds, we needed a sample period that could provide a perspective on the behaviour of fund flows as a benchmark. Since the period of the post-covid phase is one year, we found it appropriate to use a sample of similar length for comparison. Both periods were also in bull markets, improving the basis of comparison.

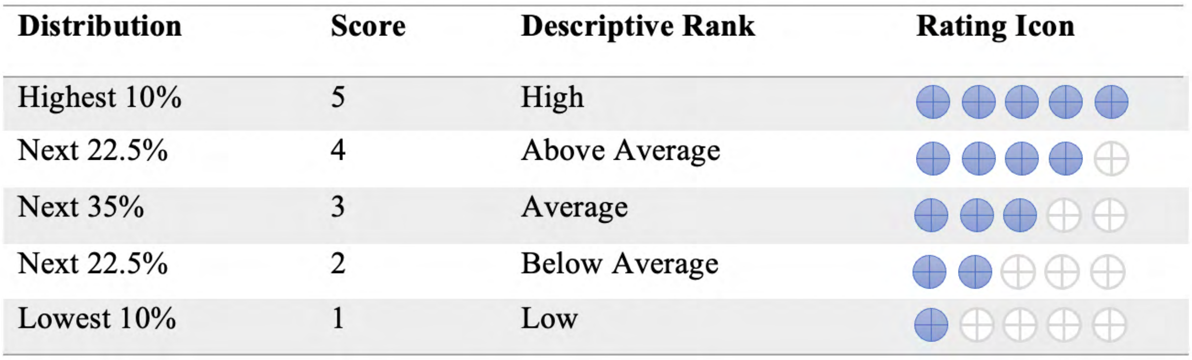
4.3 Sustainability Scores

Sustainability score is this thesis's primary variable to explain fund flow differences. In 2016 Morningstar launched the Morningstar Sustainability Rating, also referred to as MSR, intending to help investors evaluate a mutual fund portfolio's environmental, social, and governance implications. It has since been updated to use Sustainalytics' ESG-risk factors and later to include country-specific risks. This is to better account for a sovereign entity's socioeconomic well-being, derived from the government's ability to manage wealth sustainably. ESG investing's role in today's market is closely described in 3.1.

Calculating the Morningstar Sustainability Rating is a five-step process. Firstly, the portfolio is analysed to identify whether the fund is eligible to be rated by sustainability. Morningstar requires that 67 percent of the holdings are exposed to ESG to be eligible. In other words, a majority of the assets must be measurable by an existing framework for either sovereign or corporate risk for the fund to be considered eligible for rating. The second step is to calculate the corporate- and sovereign portfolio sustainability. Followed by step three, where historical corporate- and sovereign portfolio sustainability is calculated.

The fourth step combines these four calculated scores for all funds within the Morningstar Global Categories. The MSR is category relative, meaning a single fund is only compared to its peers within the global category bounds. It is also required that at least 30 portfolios within a global category are eligible for rating by either historical corporate- or historical sovereign rating. This is to ensure a broad basis of comparison for all funds. In the last step, the final rating for a fund is calculated by combining the corporate- and sovereign portfolio sustainability relative to the global category. The score is then rounded to the nearest whole number corresponding to a given number of "globes".

Ultimately, the Morningstar Sustainability rating is a ladder, where the company is given a score based on the number of globes ranging between 1 and 5. To clarify terminology in this paper regarding ratings, a fund with a score of 5 is considered a High sustainability-rated mutual fund with low ESG risk exposure, and vice versa. It is displayed by Morningstar as seen in Figure 4.1.



Distribution	Score	Descriptive Rank	Rating Icon
Highest 10%	5	High	
Next 22.5%	4	Above Average	
Next 35%	3	Average	
Next 22.5%	2	Below Average	
Lowest 10%	1	Low	

Figure 4.1: Globe system
The figure displays the globe system and equivalent measures.

4.4 Main Variable Definitions

In addition to the Morningstar Sustainability Ratings, several variables are important when analysing differences within fund flow.

4.4.1 Fund Flow

Fund flow, also referred to as net flow, is a measurement that considers the monthly net growth within each fund. The Morningstar database does not include a premade variable capturing net fund flow, they must be calculated for each fund. Following Sirri & Tufano (1998), Fund Flow can be calculated using the following formula:

$$FLOW_{ict} = \frac{TNA_{ict} - TNA_{ict-1} * (1 + R_{ict})}{TNA_{ict-1}} \tag{4.1}$$

where the TNA_{ict} is total net assets in Euros for a given fund i in country c, at time t, and $R_{ict}$is the return of that particular fund. This method is used as we do not have any data for the outflow and inflow for each fund. However, this method is used in several published articles, and we argue that the calculation is a good indicator of fund flow.

4.4.2 Monthly Return

The monthly return is the return on the fund's portfolio each period. Morningstar's calculation is based on monthly net asset value (NAV), reinvesting all income and capital-gains distributions during that month, divided by the starting NAV (Morningstar, 2022).

The variable is included to control for a relationship between net fund flow and the performance of the mutual funds. Monthly returns are presented as numerical values for the purpose of the analysis. There was also a possibility to download gross returns. However, when investigating the two variables, we found the data for gross monthly returns in comparison to monthly returns to be lacking data.

4.5 Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Net flow	179,160	0.006	0.113	−0.524	1.468
Return	179,160	0.018	0.045	−0.238	0.193
Rating	179,160	3.136	1.107	1	5
Lagged rating	177,512	3.133	1.108	1	5
Lagged return	177,512	0.017	0.055	−0.238	0.193
Post-covid	179,160	0.483	0.500	0	1
Pre-covid	179,160	0.517	0.500	0	1

Table 4.1: Summary Statistic – Whole sample period
Summarises the 179,160 observations in total. The table displays summary statistics for the entire sample period from February 2019 to April 2021. The variables Net flow and Return are displayed as numerical values. Rating is expressed through a stochastic whole-number scale from 1-5. The dummy variables Post-covid and Pre-covid are displayed as whole numbers, either 1 or 0. A value of 1 indicates that a given observation belongs in the post-covid subsample. Obs. is an abbreviation of Observations. This column describes the number of data points, meaning the sum of observations each of the variables has. The mean is the average value of all observations within the variables, and St. Dev. is the standard deviation of the variable values. Min displays the smallest value observed, and Max displays the largest value observed for each variable.

In the table, we examine the summary statistics for the entire estimation period to get a brief overview of the data. Looking at the Obs. column, there are many observations (179,160), and the distribution between before and after covid is relatively even. The Mean column is primarily interesting for the three continuous variables Net flow, Return, and Rating. The average net flow over the entire estimation period is 0.006 (0.6%), with a standard deviation of 0.113 (11.3%). At the same time, the average return is 0.018 (1.8%) with a standard deviation of 0.045 (4.5%). The lowest value recorded for net flow is negative 0.524 (-52.4%), and the largest is 1.468 (146.8%). The Return variable's most minor observation is negative 0.238 (-23.8%), and the largest observation is 0.193 (19.3%). The sample is thoroughly cleaned according to the data processing chapter to remove outliers and non-existent values before they are presented in this chapter. Even after

Winsorizing, these extreme values have been interpreted as natural extremities for the sake of this analysis.

Table 4.2: Pre-covid Summary Statistics

Panel A: Variable Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Net flow	92,642	0.002	0.111	−0.524	1.468
Return	92,642	0.006	0.040	−0.238	0.193
Rating	92,642	3.107	1.116	1	5
Lagged rating	91,724	3.101	1.117	1	5
Lagged return	91,724	0.017	0.036	−0.205	0.193

Panel B: Rating Summary Statistics

Rating	Obs.	Funds	Obs. Freq	Fund. Freq	Avg.Flow	Median	St.Dev
1	7,857	1,249	0.085	0.089	-0.003	-0.005	0.103
2	18,823	3,010	0.203	0.214	-0.002	-0.004	0.104
3	32,336	4,681	0.349	0.332	0.001	-0.003	0.107
4	22,775	3,537	0.246	0.251	0.005	-0.002	0.116
5	10,851	1,618	0.117	0.115	0.011	-0.0003	0.127

Summarises 92,642 observations before March 2020. The table displays summary statistics for the pre-covid sample period from February 2019 to February 2020. Panel A displays the variables in the sample. The variable's Net flow and Lagged return are expressed numerically. The Rating is displayed as a stochastic whole-number scale from 1-5. Panel B displays observations per rating, number of funds and the average flow.

Table 4.2 summarises the sample period from February 2019 to February 2020. There are 92,642 observations in this sample, with an average net flow of 0.002 (0.2%), an average return of 0.006 (0.6%), and an average sustainability rating of 3.107. This means the average observation has both positive net flow and return, and that there are on average more highly-rated observations than lower-rated observations.

Breaking down the observations into the different Morningstar Sustainability Rating categories, we see that most observations (34.9%) belong in the Average category. This corresponds to 4,681 unique funds with a rating of 3 (Average) at least once during the sample period. Above average and High rated categories contain 3,537 and 1,618 funds, respectively, with an average net flow of 0.005 (0.5%) and 0.011 (1.1%). Compared to the Low and Below average rated funds, with an average net flow of -0.003 (-0.3%) and -0.002 (-0.2%), there seems to be a trend that higher-rated funds have higher net flows on average in the subsample.

Table 4.3: Post-covid Summary Statistics

Panel A: Variable Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Net flow	86,518	0.010	0.115	−0.524	1.468
Return	86,518	0.031	0.047	−0.238	0.193
Rating	86,518	3.166	1.097	1	5
Lagged rating	85,788	3.166	1.097	1	5
Lagged return	85,788	0.017	0.070	−0.238	0.193

Panel B: Rating Summary Statistics

Rating	Obs.	Funds	Obs. Freq	Fund. Freq	Avg.Flow	Median	St.Dev
1	6, 380	803	0.074	0.072	-0.001	-0.003	0.094
2	16, 508	2, 223	0.191	0.199	0.005	-0.001	0.109
3	30, 374	3, 898	0.351	0.350	0.008	-0.0004	0.110
4	22, 846	2, 933	0.264	0.263	0.013	0.00000	0.116
5	10, 410	1, 293	0.120	0.116	0.023	0.002	0.138

Summarises 86,518 observations after March 2020. The table displays summary statistics for the post-covid sample period from April 2020 to April 2021. Panel A displays the variables in the sample. The variable's Net flow and Lagged return are expressed numerically. The Rating is displayed as a stochastic whole-number scale from 1-5. Panel B displays observations per rating, number of funds and the average flow.

Table 4.3 gives an overview of the sample period from April 2020 to April 2021, where there are 86,518 observations. The average net flow for this subsample is 0.010 (1.0%), with an average return of 0.031 (3.1%). The average sustainability rating is 3.166. This means the average observation has both positive net flow and return, and that there are, on average more highly-rated observations than lower-rated observations. When comparing Table 4.2 to Table 4.3, there is a trend that net flow, return, and rating means are all positively more prominent in the second period.

Displaying the observations broken down into the different Morningstar Sustainability Rating categories, the largest category is Average-rated, with 35.1 percent of the observations and 3,898 unique funds. Above average- and High-rated categories represent 26.4 percent and 12.0 percent of the observations, with 2,933 and 1,293 unique funds in each category. The average net flow in these two categories is 0.013 (1.3%) and 0.023 (2.3%). Low and Below average-rated categories have net flows of -0.001 (-0.1%) and 0.005

(0.5%). The average fund flow is higher across all ratings compared to the coefficients from Table 4.2. Furthermore, there is a proportional increase in fund flow throughout both subsamples from Low to High.

The relative frequency of observations and unique funds can provide information on whether funds in each category are longer or shorter-lived than others. A higher observation frequency than the relative unique fund frequency would indicate that fewer funds have more observations. Hence, relatively longer-lived funds for the MSR category in the sample period. This is the case for the High- and Average rated funds in the pre-covid sub-sample and for all but the Below Average rated category in the post-covid subsample. There are no clear trends when looking at the longevity of funds, nor will this be researched further in this thesis.

Country	Code	Obs.	Funds	Avg.Flow	Avg.Rating	Median	St.Dev	Freq.
Austria	AT	4,524	218	0.010	3.185	0.001	0.087	0.024
Belgium	BE	1,939	100	0.012	3.371	0.003	0.069	0.011
Switzerland	CH	4,605	239	0.002	2.634	-0.001	0.113	0.027
Germany	DE	7,255	353	0.008	3.151	-0.0003	0.088	0.040
Denmark	DK	6,237	298	0.005	3.187	-0.0001	0.119	0.033
Spain	ES	8,995	416	0.009	2.783	-0.003	0.121	0.047
Finland	FI	4,735	219	0.006	3.322	-0.001	0.112	0.025
France	FR	21,509	987	0.003	3.401	-0.002	0.097	0.111
Great Brittan	GB	21,179	960	0.002	3.112	-0.004	0.097	0.108
Ireland	IE	19,341	1,041	0.010	3.071	-0.0001	0.133	0.117
Italy	IT	2,282	110	0.003	2.856	-0.006	0.088	0.012
Lichtenstein	LI	1,877	109	-0.0003	2.904	-0.0004	0.075	0.012
Luxembourg	LU	60,958	3,220	0.006	3.092	-0.002	0.121	0.361
Netherlands	NL	3,143	162	0.0004	3.380	-0.003	0.116	0.018
Norway	NO	3,310	150	0.009	3.070	0.00005	0.088	0.017
Portugal	PT	996	45	0.003	3.189	-0.002	0.084	0.005
Sweden	SE	6,275	293	0.011	3.565	-0.0005	0.124	0.033

Table 4.4: Summary Statistics by Country
The table displays summary statistics for all observations in the sample distributed by country. Funds are unique funds from each country and Obs. is the number of individual observations. Avg. Flow is the average monthly fund flow. Avg.Rating is the average rating within each country. Freq. is the relative frequency of unique funds.

Table 4.4 displays the country-specific statistics. Even though many countries were omitted from this study to fit our scope, there are still significant differences in the included countries. Luxembourg alone represents 36 percent of all the unique funds, which is more than the next three, France, Great Brittan, and Ireland combined. Furthermore, we notice that average rating is mostly distributed around 3, ranging from 2.63 (CH) to 3.56 (SE). Besides the relative amount of funds registered in each country, there are no clear patterns of positive or negative nature concerning regionality.

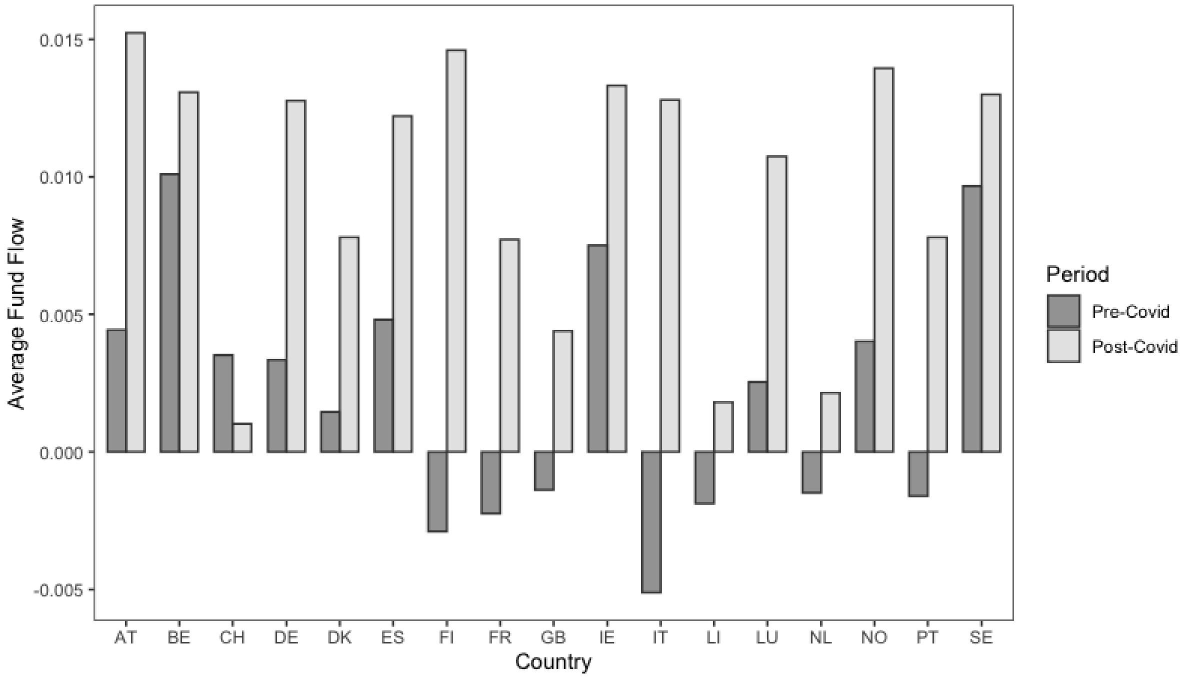


Figure 4.2: Average Fund Flow by country pre- and post-covid
The figure displays the average net fund flow numerically from the pre-covid subsample, displayed with dark grey, and from the post-covid subsample, displayed with light grey for each individual country

Based on Figure 4.2 , average mutual fund investments have increased for all countries except for Switzerland (CH) when comparing the pre-covid subsample to the post-covid subsample. This aligns with Figure A1.1, which visualises the distribution of fund flow in both sub-samples. It presents a higher concentration of observations towards greater fund flow in the post-covid sample, corresponding with the overall impression that all summary statistics provide.

5 Methodology

In this thesis, the methodology is based on several models using pooled OLS, fixed effects, and difference-in-difference estimation. The data is divided into two subsamples to explore the changes comparatively. Results from the two subsamples will be analysed to uncover whether there are any changes in market dynamics as the covid-recovery set in.

5.1 Linear regression analysis

In this section, the methodology for each linear regression will be presented, in addition to a more detailed description of how we utilise different variables and estimation tools to answer our research questions. The empirical data for fund flow changes in the sustainability-rated fund market is analysed using several linear regression methods. The models are optimised according to statistical theory by adding variables and clustering.

5.1.1 Analysis with Morningstar Sustainability Rating

To examine the effects of sustainability on mutual funds, we found the most appropriate approach to be fund flow based on sustainability-ratings. Using fund flow as a dependent variable when studying the effect of sustainability is in line with the study conducted by Pástor and Vorsatz (2020).

According to Amman et al. (2018), the ideal way to examine this kind of panel data is by comparing funds with a published sustainability rating with comparable funds with equal underlying ESG risk but no public rating. However, no existing funds satisfy this condition, and we must compare the different pools for each rating against each other. Due to the data structure and properties, we use panel regression as our primary analysis method. The research question is, therefore, partly to determine if investors value sustainability by looking at differences in fund flows across the Morningstar Sustainability Rating, which panel regression may reveal. We have converted the Morningstar Sustainability Ratings into categorical variables through integers ranging from 1 to 5, enabling a linear discrete approach to fund flow. However, when describing the non-linear context of rating on fund flow, we dummy-code or factorise the ratings to generate a more detailed view of differences between each rating. Creating categorical variables and having a non-linear perspective aligns with Amman et al. (2018), as a linear effect is not expected. Therefore, in line with Pástor & Vorsatz (2020) and Ferriani & Natoli (2020), we utilise one of the ratings as the benchmark dummy variable when exploring the non-linear effect of Morningstar Sustainability Rating on Fund Flow.

5.1.2 Linear model

To explore the linear relationship between fund flow and the explanatory variables, we will conduct a simple regression using the linear model inspired by Angrist & Pischke (2008):

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 d_{2i} + \beta_3 d_{3i} + \beta_4 d_{4i} + \beta_5 d_{5i} + \epsilon_i \tag{5.1}$$

Where y_i denotes the dependent variable fund flow, x_i the lagged return and $d_{\tau i}$ denotes the ratings as dummy-coded variables. Low-rated funds, given by d_{1i}, are integrated into the constant, whilst the rest of the ratings produce their own coefficient. For instance, d_{5i} equals 1 for High-rated funds and 0 otherwise. Due to ratings being published late each month, we found it natural to utilse the lagged rating, which is in line with previous studies.

The regressions produce data points representing observations of the explanatory variables and fund flow for the given sample period, indicating the linear relationship. However, the model does not account for omitted unobservable factors that differ for ratings but are assumed to be constant over time.

5.1.3 Fixed Effects

The panel data described in chapter 4 lays the foundation for the regressions. Panel data is cross-sectional time-series data, which refers to the pooling of cross-sectional observations of countries, funds, companies, etc., across time (Baltagi, 2005). It differs from cross-sectional and regular time-series regressions as it has two subscripts on its variables, which in this case is the security (fund) as the cross-section and months denoting the time series, given by:

$$y_{it} = \alpha + X'_{it}\beta + \epsilon_{it} \tag{5.2}$$

The error component for the disturbances is given by:

$$\epsilon_{it} = \mu_i + v_{it} \tag{5.3}$$

Which denotes unobservable individual-specific effects u_i and the remaining disturbance v_{it}. This means we can capture fund-specific effects through the ϵ_{it}

We are using several fixed effects linear models in our analysis, which is used to estimate the intrinsic characteristics of individuals. Several fixed effects linear model regressions have been conducted to examine our research question. When accounting for the fixed effects, the unobservable individual-specific effects are assumed to be fixed parameters to be estimated, while the remaining disturbances are independent and distributed identically. According to Baltagi (2005), it is an appropriate approach when focusing on, for example, firms, states, or countries. Hence, we use fixed effects at fund level in our analysis.

Furthermore, the fixed effect method holds a constant (fixed) average effect for each dummy. By including this, we can control for average differences in funds across ratings for either observable or unobservable predictors. The fixed effects eliminate the across-group variation, and we are left with within-group variations (Kellogg, u.d.). This ultimately means that the fixed effects model can prevent omitted variable bias for variables that are constant over time, resulting in a reduced chance of the model being biased as we apply the fixed effects regression to our panel data (Github, 2017). By subtracting the average over time and utilising the restriction that unobserved individual effects are eliminated ($\sum_{I=1}^{N} u_i = 0$), the fixed effect formula is given by (Inspired by Angrist & Pischke, 2008):

$$y_{it} - \overline{y_i} = \rho(D_{it} - \overline{D_i}) + (X_{it} - \overline{X_i})\delta + (\epsilon_{it} - \overline{\epsilon_i}) \tag{5.4}$$

Where $y_{it} - \overline{y_i}$ denotes fund flow, $\rho(D_{it} - \overline{D_i})$ denotes a set of dummy variables for each rating-specific intercept, $(X_{it} - \overline{X_i})\delta$ denotes the explanatory variable lagged return and ϵ is the error term. The basis-variable of the dummies are Low-rated funds. Also, i denotes fund, while t represents the period.

We have divided the analysis using fixed effects into different regressions based on the entire estimation period and sub-sample periods. This is to enable an understanding of how the variables behave over time and present the degree of impact ratings has on fund flow both prior to and post to March 2020. Furthermore, lagged return is included across all regressions to improve the model goodness of fit and reduce omitted variable bias. To account for heteroscedasticity and un-modelled dependence among the error terms, it is common to cluster the variables for fixed effects estimation in panel data, thereby creating cluster-robust standard errors (Pustejovsky & Tipton, 2017). The fixed effects models presented in empirical results are not clustered on fund level to add cluster-robust standard errors. However, the appendix contains the same table where clustering is included (Table A1.1).

5.2 Difference-in-difference regression

The pre- and post-covid panels from the linear – and fixed effects method is used to examine patterns of behaviour for the ratings across time within the sample periods. However, it does not state the direct fund flow differences between the two periods. To control for this, utilising an interaction term could be beneficial to improve the understanding of general fund flow across time and enable specific testing to provide further depth to our analysis. Our approach of choice is a difference-in-difference method.

The difference-in-difference (DiD) estimator is a popular tool when evaluating a treatment of interest between two periods, in our case, the covid-19 recession (Abadie, 2003). The primary purpose of this regression is to determine if there are significant changes in fund flow for High and Low-rated funds when examining the pre-intervention phase and the post-intervention that follows after the treatment period, in this case, the post-covid subsample. Therefore, a panel only consisting of funds with a lagged rating of either 1 or 5 is included, testing if High- and Low-rated funds impact fund flow differently across time. A dummy that separates the time periods must be established to create the DiD-estimator. In this case, the time variable is created, equal to one if the date is in the post-covid phase and zero otherwise. The next step is to create the dummy variable treatment that identifies the group, meaning High-rated funds with a rating value of 5 is equal to one, while Low-rated funds equal 0. As both variables are produced, the DiD-estimator is ready, consisting of an interaction between the treatment and time variables. Following Angrist & Pischke (2008), we get:

$$Y_{ist} = \alpha + \gamma T_s + \lambda d_t + \beta(T_s * d_t) + \epsilon_{ist} \tag{5.5}$$

Where α ,γ ,λ , β are unknown parameters and ϵ_{ist} is the random unobserved error term containing determinants that the model omits. Also, i denotes fund, s denotes rating (High or low) and t denotes the period (Pre- and post-covid). T_s represents a dummy for the rating variable, d_t the time-dummy, and $(T_s * d_t)$ the DiD-estimator.

5.3 Country-specific effects model

As an additional perspective, we want to explore how fund flows are affected by sustainability ratings within each country. The approach for this analysis includes creating a difference-in-difference regression for each individual country without fixed effects and a linear regression factorised by rating. Similar to the last chapter, we only explore High-rated versus Low-rated funds. The output will include two scatterplots based on Table A1.2 which can be found in the appendix. The first plot shows whether High-rated funds have significantly higher fund flow than Low-rated funds for the whole period to add a general perspective on sustainable fund investment within all countries. The second plot displays the difference-in-difference estimator. By doing so, we can examine how mutual fund flow within countries behave in line with sustainability-rating over time, exploring demographic differences. Lastly, the logic and formula are equivalent to the presented material in section 5.2.

6 Empirical Results

Building on the methodology, the empirical results of our analysis will be presented and visualised in the following chapter. This includes analysing the entire sample period, the pre- and post-covid phases, and country-related perspectives. The analysis contains plots based on average statistics, several linear and fixed effects regressions, and a difference-in-difference approach.

6.1 Analysis of fund flow to ratings

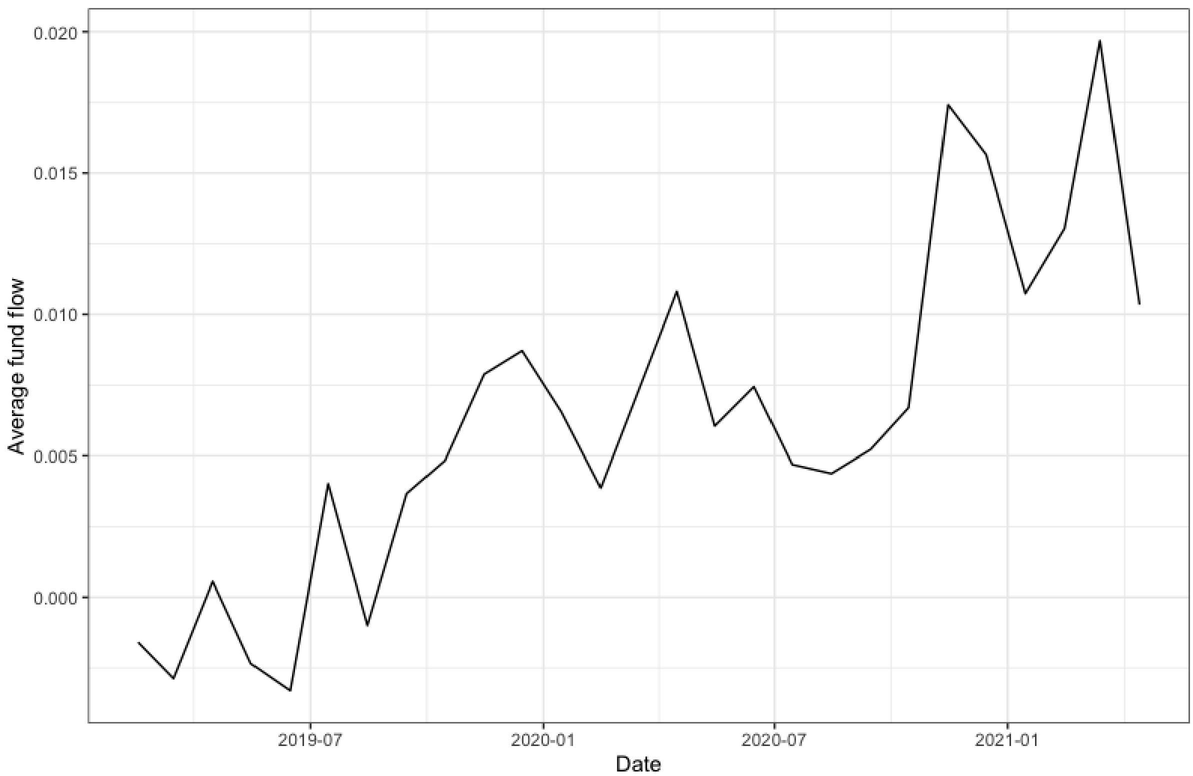


Figure 6.1: Average Net Fund Flow for all rated funds
The plot displays the average monthly net fund flow, for all funds, over the entire sample period. March 2020 is not included.

Figure 6.1 displays the average fund flow for all funds with a Morningstar Sustainability Rating in the total sample, excluding March 2020. Average net flow has experienced an increase within this timeframe, starting at -0.15 percent in February 2019 and reaching 1.97 percent by April 2021. However, there was a steep decline from February 2020 to April 2020 due to the Covid pandemic. The lowest fund flow recorded during the full period was -0.3 percent, which is a significant negative outflow considering the size of these funds combined. Further analysing the figure, there is a clear trend of increasing net fund flow throughout the sample period.

Figure 6.2: Average Fund Flow by Morningstar Sustainability Rating
The figure displays the monthly average net fund flow numerically, grouped by Morningstar Sustainability Rating.

In Figure 6.2, the average fund flow is categorised by sustainability rating. We notice that all ratings have increased on average over the whole period. Although the difference between some ratings are unclear, higher-rated funds have experienced higher net fund flow than lower-rated funds, which is especially evident in 2020. This is further supported by Figure A1.2, providing the average fund flow for each rating.

	Dependent variable:						
	Net Flow						
	OLS			*felm*			
	(1)	(2)	(3)	(4)	(5)	(6)	(6b)
lagrating	0.003***	0.005***	0.004***	0.001	0.002	0.001**	0.001*
	(0.0003)	(0.0003)	(0.0002)	(0.001)	(0.001)	(0.0004)	(0.001)
lagreturn	0.070***	0.033***	0.041***	0.045***	0.006	0.017***	0.017***
	(0.010)	(0.005)	(0.005)	(0.009)	(0.005)	(0.004)	(0.004)
Constant	−0.010***	−0.008***	−0.009***				
	(0.001)	(0.001)	(0.001)				
Fixed effects	No	No	No	Yes	Yes	Yes	Yes
Cluster	No	No	No	No	No	No	Yes
Observations	91,724	85,788	177,512	91,724	85,788	177,512	177,512
R^2	0.002	0.003	0.002	0.206	0.237	0.160	0.160
Adjusted R^2	0.002	0.003	0.002	0.131	0.157	0.116	0.116
Note:					*p<0.1; **p<0.05; ***p<0.01		

Table 6.1: Fund flows in Response to MSR
Describes the pooled OLS and fixed effects regression with fund flow as the dependent variable and the discrete lagged rating and lagged return as explanatory variables. Column (1), (2) and (3) shows the pooled OLS model, while (4), (5) and (6a) displays the Fixed Effects model for respectively pre-covid, post-covid and the entire sample period. Column (6b) is the same regressions as in (6a). However, it includes cluster-robust standard errors on fund level.

Running the linear approach with lagged rating, we notice that the pooled OLS suggests that for every increase in rating, the expected average fund flow increases by 0.4 percent across the whole sample period. This is significant at the 0.01 level across all three models. It's worth mentioning that the post-covid sample provides a higher coefficient, suggesting that ratings have more impact on fund flow during this phase.

The fixed effects model also produces higher coefficients in the post-covid phase. However, they are not significant, meaning we cannot conclude that there are any differences between the two periods when utilising the linear-rating approach. The whole estimation period is significant in both the non-clustered and clustered models, indicating a linear relationship. It suggests that monthly fund flow increases with 0.1 percent per increase in rating, meaning High-rated funds have a lagged rating coefficient of 0.5 percent (0.001*5).

	Dependent variable:					
	Net Flow					
	OLS			*felm*		
	(1)	(2)	(3)	(4)	(5)	(6)
Below average	0.001	0.006***	0.004***	0.002	0.007*	0.004**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.004)	(0.001)
Average	0.004***	0.009***	0.006***	0.002	0.006	0.004***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.004)	(0.002)
Above average	0.007***	0.013***	0.010***	0.003	0.007	0.005***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.005)	(0.002)
High	0.013***	0.023***	0.018***	0.004	0.012**	0.005**
	(0.002)	(0.002)	(0.001)	(0.003)	(0.005)	(0.002)
Constant	−0.004***	−0.002	−0.003***			
	(0.001)	(0.001)	(0.001)			
Fixed effects	No	No	No	Yes	Yes	Yes
Observations	91,724	85,788	177,512	91,724	85,788	177,512
R^2	0.001	0.003	0.002	0.206	0.237	0.160
Adjusted R^2	0.001	0.003	0.002	0.130	0.157	0.116
Note:				*p<0.1; **p<0.05; ***p<0.01		

Table 6.2: Fund Flows in response to MSR (factorised)
The table displays the pooled OLS model and the fixed effects linear model of the panel data with factorised ratings. Hence, describing the non-linear relationship across lagged ratings. Columns (1) and (4) show the regression for the pre-covid subsample, (2) and (5) for the post-covid subsample and (3) and (6) for the whole time period. The OLS regressions are based on Formula 5.1, while the felm regressions are based on Formula 5.4 . The dependent variable is Fund flow presented numerically, with 177,512 observations in total. The explanatory variables are the factorised lagged rating (t-1) with Low-rated funds as the base-dummy and the lagged return (t-1). The felm regressions are not clustered in this table. See Table A1.1 in the Appendix.

Table 6.2 describes our findings when analysing the different time periods using several panel regression methods. Column (1), (2) and (3) is based on the linear model approach. In column (1), we find a significant relationship below the 0.01 level for all ratings except Below average. This suggest that there is no significant difference between fund flow to Low and Below average rated funds in the pre-covid phase, which can be interpreted through the constant as negative 0.5 percent. Average rated funds increase expected fund flow by 0.004. However, still negative expected fund flow at -0.01 (-0.005 + 0.004). Above average and High-rated funds increase the expected average fund flow by 0.7 percent and 1.2 percent, respectively. The post-covid phase in column (2) shows a similar pattern. However, the coefficients are generally greater, meaning that fund flow is higher across all ratings. For instance, the constant indicates that we have insufficient evidence that Low-rated funds differ from zero and that High-rated funds have a 2.3 percent higher fund flow. We also notice that for higher ratings, the coefficient increases. Furthermore, we add column (3) to see how ratings behave over the entire estimation period. This confirms that all ratings have experienced different fund flows throughout the estimation period, with highly rated funds increasing the most. The R-squared suggests that these models capture between 1.0 and 3.0 percent of the variations in the data.

Columns (4), (5) and (6) represent the fixed effects linear model that is used for the panel regression. Column (4), the pre-covid phase, contains the lagged rating and return as explanatory variables. In this regression, no significant variables are produced, meaning that we cannot, with confidence, suggest that lagged rating has affected monthly fund flow during the year before March 2020. The results from the pre-covid phase, presented in column (5), state that there are some significant coefficients. In particular, High-rated funds below the 0.05 level and Below average-rated funds below the 0.1 level. High-rated funds have an expected monthly fund flow of 1.2 percent higher than Low-rated funds. Considering the whole estimation period in column (6), the results suggest that the Morningstar Sustainability Ratings affect fund flow, where all ratings are significantly greater than Low-rated funds. The R-squared ranges between 0.16 and 0.24.

We also added the lagged returns to get a perspective on the flow-performance relationship for sustainability-rated funds. The linear model suggests that an increase in lagged return increases average fund flow, meaning that returns are considered a positive attribute for sustainability-rated funds. In the fixed effects model, we notice that lagged return in the post-covid phase is not significant, while being significantly positive during the pre-covid phase.

6.2 Difference-In-Difference analysis

	Dependent variable:			
	Net Flow			
	OLS		*felm*	
	(1)	(2)	(3)	(4)
high	0.013***	0.012***	−0.002	−0.002
	(0.002)	(0.002)	(0.008)	(0.008)
post_cov	0.002	0.002	−0.004*	−0.004*
	(0.002)	(0.002)	(0.002)	(0.002)
did_est	0.010***	0.010***	−0.0001	−0.0001
	(0.002)	(0.002)	(0.003)	(0.003)
lagreturn		0.045***		0.013
		(0.011)		(0.010)
Constant	−0.004***	−0.005***		
	(0.001)	(0.001)		
Fixed effects	No	No	Yes	Yes
Cluster	No	No	No	No
Observations	35,219	35,219	35,219	35,219
R^2	0.008	0.008	0.211	0.211
Adjusted R^2	0.008	0.008	0.127	0.127
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 6.3: Difference-in-Difference with High-rated funds as treatment
*The table displays the difference-in-difference regression and is based on Formula 5.5. Columns (1) and (2) display the OLS results, while columns (3) and (4) are fixed effect regressions. Pre-covid is the pre-intervention period, while post-covid is the post-intervention period. The "treatment" group is High-rated funds. The Constant represents Low-rated funds pre-covid. The DiD-estimator represents treatment*time (High*Post-covid). The dependent variable is fund flow.*

Column (1) is the DiD-analysis regressed as an OLS model without including additional control variables. The Constant term represents the untreated group, with a coefficient of negative 0.004. Meaning the net fund flow to Low-rated funds is expected to be negative on average. High-rated funds are expected to have a positive net flow of 0.009 (-0.004+0.013) pre-covid. The coefficient for post-covid in columns (1) and (2) is 0.002, yet not significantly different from zero. Hence, whether Low-rated funds experienced an increased net flow after March 2020 is unclear. The difference-in-difference estimator, which explains the treatment effect, has a coefficient of 0.010 and is statistically significant at the 0.01 level. The impact of High ratings after March 2020 is thereby a 1.0 percent increased net fund flow.

The results of lagged return as an additional control variable for OLS can be seen in column (2). The constant is -0.005, suggesting potential omitted variable bias in the column (1) regression. The expected fund flow for Low-rated funds is -0.50 percent. Funds rated High have an expected fund flow of 0.007 (-0.005+0.012), which is smaller than the previous regression showed. The lagged return has a coefficient of 0.045, suggesting that the return can explain some variations in fund flow. Looking at the model's explanatory power, it does not seem like including lagged returns increases the model's ability to capture the variation of net fund flow.

Columns (3) and (4) display the results from identical model specifications as columns (1) and (2), but now controlling for fixed effects within each fund. The only significant variable for both columns is the post-covid dummy with equal coefficients of -0.004. All other variables are not significantly different from zero. The fixed effects model does not capture the same level of significance that OLS provide. This could indicate that within variation in funds are lower opposed to the OLS assumption that model parameters are common across funds and time.

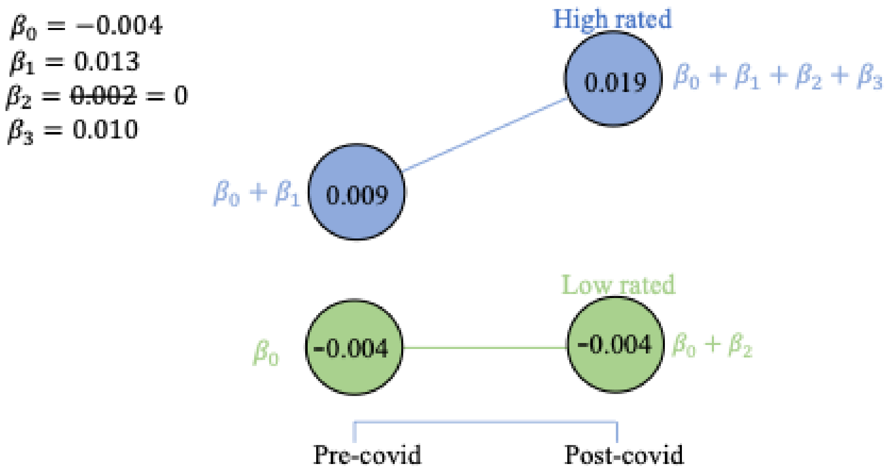


Figure 6.3: Difference-in-Difference coefficients visualized
The figure is based on Table 6.3 column (1). It visualises the effect of High and Low ratings in the difference-in-difference model. The constant β_0 is the average outcome of low-rated funds pre-covid. The rating β_1 coefficient represents the difference between the groups before the treatment. Post-covid β_2 represent the trend over time, meaning how general fund flow has changed in the post-covid phase. It is not significant, therefore set to zero. The difference-in-difference estimator β_3 represent gains for High-rated funds in the post-covid phase compared to pre-covid.

The difference-in-difference estimate is (0.019 – 0.009) – (-0.004 – (-0.004)) = 0.01, which equals the difference-in-difference estimator from Table 6.3.

6.3 Fund Flow for Sustainability rated funds by Country

To break down demographic differences in line with the exploratory research question (3), this section presents empirical results for fund flow to High versus Low-rated funds within each country. In addition, it provides a perspective utilising the difference-in-difference approach on whether High-rated funds have increased fund flow across time considering regions.

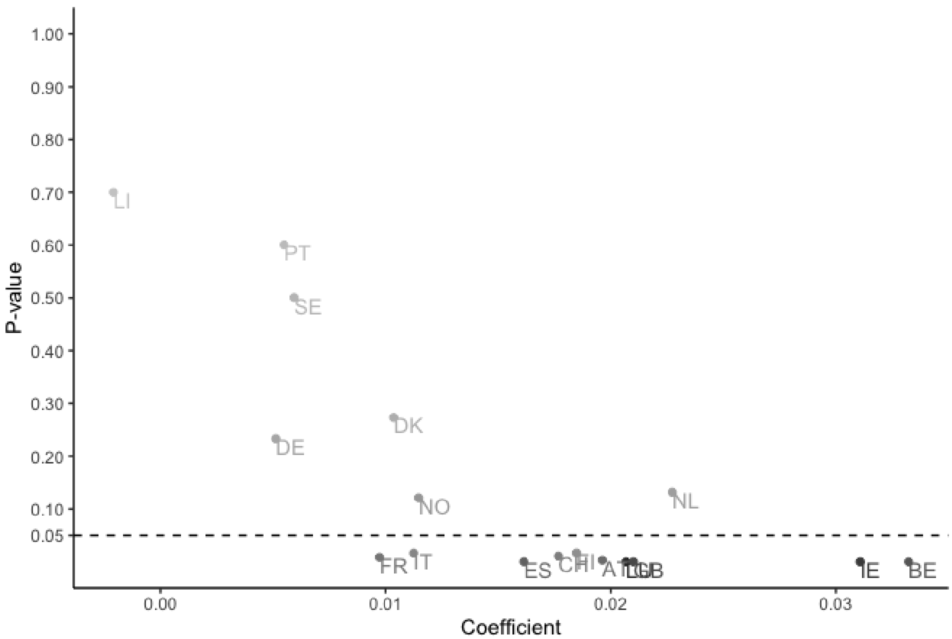


Figure 6.4: Scatterplot on High versus Low ratings by country
The figure is a plot of regression coefficients for each country. The X-axis represents the coefficient of High-rated funds as a dummy variable, while the Y-axis represent the p-value of each coefficient. Only High- and Low-rated funds are included in the underlying data, meaning the coefficient compares the two different ratings. The time frame is the entire estimation period from February 2019 to April 2021. A respective table can be found in the appendix providing the extended results in table format.

Figure 6.4 describes our findings when comparing High-rated funds to Low-rated funds within each country for the entire timeframe. All coefficients are positive except LI, meaning almost all countries had excess fund flow to High-rated funds compared to Low-rated funds in the same domicile. The p-values disclose that 10 of the 17 included countries have statistically significant coefficients at the 0.05 level. Belgium (BE) had the highest coefficient of 0.033, meaning High-rated funds had higher expected net flow of 3.3 percent compared to Low-rated funds.

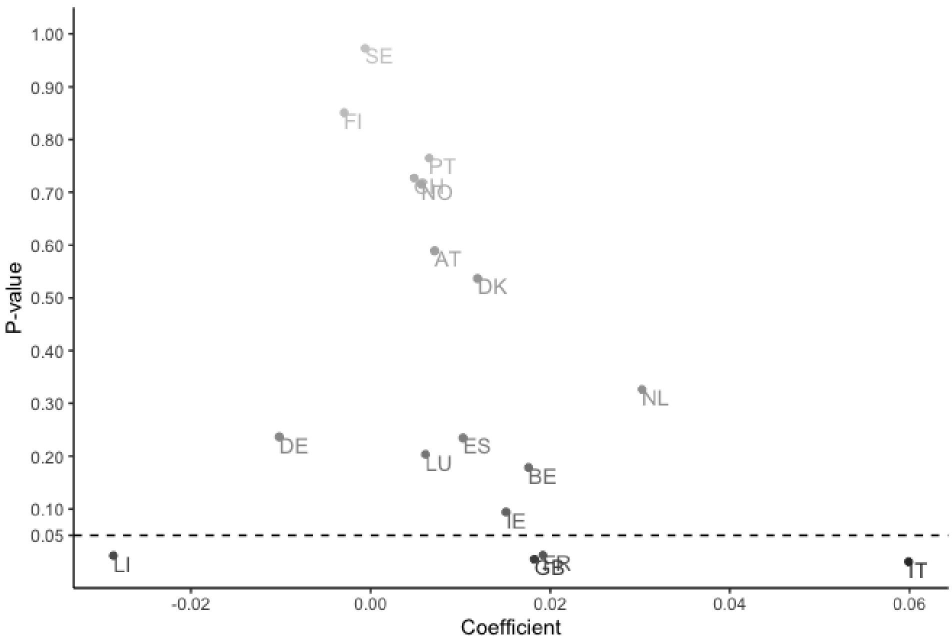


Figure 6.5: Scatterplot with difference-in-difference estimator by country
The figure is based on the difference-in-difference regressions for each country. The X-axis represents the difference-in-difference estimator, and the Y-axis represent the p-value of the coefficient. It contains the entire sample period, using the pre-covid phase as the pre-intervention and post-covid as the post-intervention. The difference-in-difference estimator is calculated using the treatment effect multiplied by a dummy variable for the time effect representing the post-covid time frame. The table can be found in the appendix (Table A1.2).

Figure 6.5 displays the difference-in-difference estimator for all included countries on the x-axis, with respective p-values plotted on the y-axis. Analysing the figure, we see that all coefficients except LI, DE, FI, and SE are positive. They indicate excess net flow to High-rated funds in the post-covid timeframe for most countries. However, only the coefficients for countries LI, GB, FR, and IT are statistically significant. In Italy (IT), the difference-in-difference estimator has a value of 0.06, meaning the High-rated funds experience a 6.0 percent excess fund flow post-covid. For Great Brittan (GB), the coefficient is 0.018, indicating an excess flow of 1.8 percent, while France (FR) had an excess flow to High-rated funds of 0.019 or 1.9 percent. Lichtenstein (LI) was the only country with a negative and significant coefficient. High-rated funds had a negative net flow of 0.029, meaning high-rated funds had 2.9 percent less inflow post-covid.

Considering there were few significant coefficients from the regressions, it is hard to draw a general conclusion for all countries included. Furthermore, the countries with significant p-values had quite varying results from the difference-in-difference analysis.

7 Discussion

In this section, we will discuss the results retrieved from the empirical results in line with our research questions, hypothesises and literature review.

7.1 Discussion of empirical results

Average statistics

Figure 6.1 displays how the average fund flow has increased over the entire period within the sustainable mutual fund market. There is an apparent increase in average fund flow throughout the plot. This looks especially applicable in the post-covid timeframe, supporting the general thought of a soaring mutual fund market. From Figure 6.2, we notice that all ratings have increased throughout the entire estimation period. This could be a result of increased focus on sustainability among the public, and that investors to a higher degree recognise ESG factors as drivers of value (McKinsey, 2017).

Flow-Rating relationship with OLS and Fixed effects

The OLS regression results showed that rating positively affected fund flow over the entire period. From Table 6.1 , an increased rating is associated with a linearly increased fund flow of 0.4 percent. A High-rated fund is therefore expected to have an increased fund flow of 1.6 percent compared to Low-rated funds. Table 6.2 presents the ratings as dummy variables to control for non-linearity across the levels. The results show minor variations between the rating levels for the entire period, suggesting a non-linear relationship. Using the fixed effects methodology, the results are similar. A higher rating correlates with the increased net flow, however, not as ample as in the OLS models.

Interpreting the full estimation period results, we see evidence supporting our hypothesis for research question (1). We expected investors to value sustainability greatly when allocating capital to mutual funds. The predictions were mainly based on previous studies conducted by Amman et al. (2018) and Hartzmark & Sussman (2019), combined with Google Trend data. Unlike our approach, these studies benchmarked a selected sample of funds before and after they received their rating. Nevertheless, their results provide a satisfying view of sustainable mutual fund flow applicable to our thesis. Furthermore, Amman et al. (2018) found significant differences between Low and High-rated mutual funds in their study of lagged ratings. Similar to the results that our analysis has produced. However, the coefficients from our analysis were smaller than expected. This is especially true for fixed effects regression results, where the linear effect was only 0.1 percent, and the dummy variables for ratings 2-5 all had either 0.4 percent or 0.5 percent increased flow. Hence, the fixed effects results were more uniform than we expected, in contrast to the OLS results, which showed greater differences between different rating levels.

When analysing the pre- and post-covid periods separately, the results suggest that the impact from ratings is different between the two periods. From the OLS regressions in Table 6.1, the post-covid subsample coefficient was 0.2 percent larger than the equivalent for the pre-covid subsample. The constant also increases, suggesting an overall increase in fund flow. Furthermore, this is increasingly visible in Table 6.2, where the growth difference for each level is visualised. The increase between the two periods is again the largest for the High-rated funds. Similar results are found when using the fixed effects models. However, only the High-rated funds have a significant coefficient of 0.012 at the 0.05 level post-covid.

How the effect of the Morningstar Sustainability Rating has developed over time was also crucial for our initial motivation. Based on studies by Pástor & Vorsatz (2020) and Ferriani & Natoli (2020), along with general market information, we hypothesised that post-covid would mark a shift towards greener investments. This is mainly due to the market climate, which allowed investors to increase their green exposure, taking advantage of the government stimulus packages to reduce initial risk. Our analysis confirms that there was a shift which can be explained in two parts. Firstly, following the Covid-19 crisis, there was a general increase in the flow to MSR-rated funds. Meaning investors invested more than previously in the mutual fund market. The second part is where the increased flow ended up. Columns (1) and (2) in Table 6.2 present evidence that more of the new flow went to higher-rated funds. The Morningstar Sustainability Rating greatly affected the flow to sustainable mutual funds following the Covid-19 market crash, while the effect was less visible earlier. This will be further discussed using results from the difference-in-difference analysis.

Considering the results presented using OLS and the fixed effects models, we argue that European investors value the Morningstar Sustainability Rating of funds. However, the magnitude might be moderate. Though, there has been a positive shift following the Covid-19 market crash. There were also increased sustainability-rated fund investments, notably favouring the highest-rated mutual funds, which is further supported and visualised in Figure 6.2.

The fixed effect models experienced much lower significance than the OLS models did. This could be because degrees of freedom are consumed to control for time-fixed effects, weakening the model's accuracy. Another explanation is that the OLS models assume each observation is independent, contrary to the fixed effect model that can analyse the same object across time, which might result in biased OLS results. Suppose funds are relatively equal in the pre- and post-covid phase. In that case, the effect of controlling for average differences in funds across ratings is lower than when we assume linearity in parameters. This could also partially explain the deviating results between OLS and the fixed effects regressions. Nevertheless, it is important to keep in mind that FE models solely identify effects based on within-individual changes, whereas pooled OLS models consider between-individual variation (Bell & Jones 2015).

Pooled OLS is more commonly employed when you select a different sample for each period of the panel data. A problem with pooled OLS models is that the presence of heteroskedasticity can render the regression inadequate. This is a common problem when working with panel data models (Saeed et al., 2018), and a Breusch Pagan test could uncover potential issues related to heteroskedasticity. Based on the models' explanatory power, and the potential biases OLS is exposed to, we believe the fixed effects model provides the most applicable results. This is in line with Collischon & Eberl (2020), who argue that FE models are preferred when time-constant unobserved heterogeneity is likely to be a problem.

When considering the flow-performance relationship, we expected returns to be less important as sustainability-oriented investors are less prone to act on poor returns (Taylor, 2020). The results suggest that investors weighted returns more heavily before the Covid-19 crash. We argue this is due to the extraordinary market situation following the recession, with fewer risk nuances and seemingly ever-growing assets.

Flow-Rating relationship with DiD

We conducted several difference-in-difference studies to further analyse the positive shift discovered by the OLS, fixed effects models, and summary statistics. These studies aimed to isolate the effect of a fund rated High or Low compared to the respective opposite. It was evident that Low-rated funds experienced negative fund flow throughout the entire period. High-rated funds had a 0.9 percent net flow pre-covid, 1.3 percentage points higher than Low-rated funds. For Low-rated funds, the results from column (1) in table 6.3 indicate no significant differences between the two periods. Interpreting the difference-in-difference estimator, we notice that High rating in the post-covid environment equals a 1.0 percent increased fund flow. The results were also checked for robustness by adding another variable to the regression, which resulted in minor alterations that did not affect the conclusion noteworthy. Subsequently, we find that investors within the sustainable mutual fund market disproportionally reallocate towards High-rated funds over time, while fund flow for Low-rated funds experienced no significant changes. This is similar to what Sirri & Tofano (1998) experienced with high and low-performing funds, where consumers were reluctant to flee low-performers while simultaneously experiencing a flock towards high-performers.

Pástor & Vorsatz (2020) and Ferriani & Natoli (2020) argued that investors favour High-rated funds, especially during the recession and after. Their results are similar to ours, supporting the hypothesis that the pandemic marks a shift in sustainable investing, and further suggesting that ESG has gained hype and worked as an equity vaccine. The shift could partly be a result of additional focus or hype through increased media attention as it, according to Sirri & Tofano (1998), has an impact on current growth within mutual funds. Additionally, this implies that Pástor & Vorsatz' (2020) view of sustainability as a necessity, and the reallocation towards High-rated funds, is relevant in a broader time perspective.

Findings from the difference-in-difference analysis outline the amplitude of the effect of a fund being rated High vs. Low. We see a positive shift from pre-covid to post-covid, aligning with our hypothesis. Although a 1.0 percent increase is small considering the relative importance of ESG, the effect is still significant because of the sheer size of the European mutual fund market. Furthermore, we have to be cautious, as it, in most cases, is incorrect to interpret coefficients from OLS and Fixed effects as causal effects rather than partial correlations. To be credible, causal effect identification tends to require an exogenous shock, as in an experiment (Collischon & Eberl, 2020).

Geographical differences in fund flow

Throughout Europe, many countries are different from one another. This presented an opportunity to explore the European mutual fund market further. Through difference-in-difference methodology, we examined the effect of a High-rating compared to a Low-rating within each country. The results were quite varying. More than half of the sample countries had a significant High-rating coefficient for the entire period. For the difference-in-difference estimator, only countries Lichtenstein, Great Brittan, France, and Italy were significant.

As more than half of the countries from Figure 6.4 have significant coefficients at the 0.05 level, there seems to be a trend that several domiciles favour High ratings over Low ratings. Furthermore, suggesting that more than half of the sample countries consider Morningstar Sustainability Rating a positive attribute when investing in the sustainable mutual fund market. This implies some differences across regions which supports our hypothesis for research question (3). We would expect this share to be higher, considering Europeans generally view themselves as "Green" (AMA staff, 2019). Additionally, there is no clear pattern between the countries' sustainability score (from Morningstar Sustainability Atlas) and fund flow to High-rated funds.

The findings from Figure 6.5 state that four of the countries had a significant difference-in-difference estimator below the 0.05 level. Meaning only a few of the domiciles experienced a higher average fund flow into High-rated funds post-covid. From these, England, France and Italy had positive coefficients. While Funds from Lichtenstein, surprisingly, had a significant negative difference-in-difference estimator of -0.029. This could mean that ESG is considered a luxury rather than a necessity, leading people to selling their assets as the panic following Covid-19 started to evolve.

Considering how well-developed most of the sample countries are regarding ESG, we would expect a higher share of significant results. Table 4.2 shows that the average fund flow for most countries was higher post-covid. However, this increase had less impact than expected on the fund flow to High- and Low rated funds. A possible explanation for these deviating results could be the small sample sizes for each country. Looking at Table 4.4, many countries have less than 5,000 observations for the entire sample, which could affect the accuracy of the results, especially the significance.

We hypothesised that there would be differences within the different domiciles and that especially western European countries with stable economies would have significantly greater fund flow to High-rated funds. Our results are scattered, with few precise difference-in-difference results or trends visible. However, some differences between the countries are still interesting to observe. From these results we can also derive that regression as a tool of analysis has apparent weaknesses. Aggregated results from a selection of European countries does not necessarily mean the conditions are applicable in each domicile separately, and must be interpreted with this in mind.

7.2 Limitations and future research

Throughout this study, many uncertainties and limitations should be considered when reading. These are mainly related to the sheer complexity of sustainability and how this complexity is processed and handled by institutions. Additionally, the study could be improved in several ways that are not convenient for us to correct at this point. Correcting this could enhance results and lay an improved foundation for future research.

We have used the Morningstar Sustainability Rating as a direct rating for the underlying sustainability of a fund. However, this is only partially accurate to what the rating actually means. It is not directly a rating of underlying sustainability. It is a relative measure of ESG risk in the portfolio compared to industry peers in the same Morningstar Global Category. This means a fund may have a better rating than others, even though their underlying ESG risk is equal. Because they belong to different peer groups, they have different standards for the rating system. In our analysis, we chose to interpret the rating system in the context and manner it would be presented to an investor. Even though this approach might not utilize the full information potential.

Another aspect that may limit the causality of the results is challenges associated with ESG score reliability, which according to Doyle (2018), is, for example, rating biases, inconsistency between score providers and lack of ability to identify risks. Environmental and Social disclosures have no standardised rules nor satisfying control mechanisms verifying the reported data. Consequently, the agencies apply subjective assumptions of the firm's ESG state when identifying risks, in addition to company size –, geographic

– and industry sector biases. Berg et al. (2019) examined ESG scores across six major providers of ESG rating data, where they found an average correlation of 0.54, ranging from 0.38 to 0.71 between the different providers. Disaggregating the different ESG dimensions, the environmental correlation equals 0.53, the social factor equals 0.24, and governance has the lowest correlation of 0.30. Due to the variations among ESG score providers, inconsistencies might limit our results to Morningstar/Sustainalytics data only. It could therefore be beneficial to investigate other agencies and perform a comparative analysis to strengthen our results. Something that might be suitable to research in the future.

Sustainability and ESG have been discussed in line with mutual fund flows throughout this thesis. However, Jesse (2022) suggests that integrating ESG factors into decision-making does not necessarily make an organisation sustainable. An organisation with a given ESG policy uses ESG frameworks in their decision-making, but whether they are sustainable depends on the trade-off between various categories and weights assigned to each dimension. Therefore, it is hard to draw any conclusions about increased fund flow to sustainability-rated funds actually having an impact on sustainability.

There are many opportunities to complement and expand our study in future research. Extending the timeline to more recent data could be interesting, considering the extreme development we have seen over the last year. Additionally, including non-rated – and index funds could provide a solid foundation to research behavioural differences between rated and non-rated funds for both active and passive funds. Another interesting question that arises is which of the three ESG factors are the most important drivers when investing in mutual funds and whether investors differentiate between them. Adding more variables to avoid omitted variable bias and increasing the complexity of other vital drivers within fund flow is also possible. This could be retail vs institutional investors, fund size & age, and more.

8 Conclusion

In this thesis, we investigate fund flow to sustainability-rated mutual funds one year before and after the Covid-19 recession in March 2020. We believed it would be interesting to analyse how the fund flow has developed and see what changes the market has been exposed to during this period. As Europe is a complex market, we also include studies for country-specific changes, aiming to observe any regional trends or differences between the sample countries included.

Regardless of the sustainability rating, we find that fund flow has increased throughout the sample period. We find evidence that the dynamics in the European market have seen a change from pre-covid to post-covid where general fund flow has increased across ratings compared to Low-rated funds, especially for High-rated funds. The OLS and fixed effects models also point to High-rated funds experiencing greater fund flow in comparison to Low-rated funds. These results indicate that ratings are treated as a positive attribute considering higher ratings accumulate more fund flow, implying that investors value sustainability levels and use the information actively. Our difference-in-difference models support these claims, where the results indicate a shift towards increased inflow post-covid for High-rated funds. This further suggests that investors within the sustainable mutual fund market have increased their investments over time.

From the country analysis, we find less evidence than anticipated. 10 out of the 17 included countries had significant results when analysing High-rated compared to Low-rated funds. All results support the general trend of increased inflow to High-rated funds. Divergent results were discovered for the difference-in-difference analysis of High-rated funds post-covid. England, France, and Italy had results suggesting increased fund flow to High-rated funds post-covid, while Lichtenstein had a decrease. We see a trend favouring High over Low ratings. However, as a general result for most countries, there is little evidence suggesting excess flow to High-rated funds post-covid.

We have applied several panel regression methodologies throughout this study to examine causal relationships between fund flow and rating within sustainable mutual funds. Based on articles, literature, and our findings, we argue that Morningstar's sustainability rating serves as a very accessible source of information that investors frequently use and value.

As ESG has established itself as a necessity over time, we present evidence of a general increase of capital fund flow to the sustainable mutual fund market, with proportionally greater inflows towards higher-rated mutual funds. Given the state of the world, both politically and economically. It will be interesting to see how sustainability-ratings will affect investment behaviour long-term and whether sustainability-rated assets continue to increase their market position when inflation, war and energy prices become severe worldwide problems.

References

Abadie, Alberto (2003). *Semiparametric Difference-in-Differences Estimators.* The Review of Economic Studies.

AMA staff (2019). *New Survey Looks at the "Greening" of Europe vs. the U.S.* American Management Assocation.

Amman, Manuel; Bauer, Christopher; Fischer, Sebastian; Müller, Philipp (2018). *The impact of the Morningstar Sustainability Rating on mutual funds flows.* European Financial Management.

Angrist, J.D. & Pischke, J. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.*

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data (Vol. 3).* Wiley Sons.

Bell, Andrew & Jones, K. (2015). *Explaining fixed effects: Random effects modeling of time-seriescross-sectional and panel data.* Political Science Research and Methods.

Berg, F.; Kölbel, J.; Rigobon, R. (2019). *Aggregate Confusion: The Divergence of ESG Ratings.* SSRN Electronic Journal.

Bialkowski, J.; Starks, L. T. (2016). *SRI funds: Investor demand, exogenous shocks and ESG profiles.*

Boffo, R.; Patalano, R. (2020). *ESG Investing; Practices, Progress and Challenges.* OECD.

Collischon, M. Eberl, A. (2020). *Let's Talk About Fixed Effects: Let's Talk About All the Good Things and the Bad Things.* Kölner Zeitschrift für Soziologie und Sozialpsychologie.

De Spiegeleer, Jan; Höcht, Stephan; Jakubowksi, Daniel; Reyners, Sofie; Schoutens, Wim (2020). *ESG: A New Dimension in Portfolio Allocation.* SSRN.

Doyle, T. M. (2018). *Ratings That Don't Rate: The Subjective World of ESG Ratings Agencies.* American Council for Capital Formation.

Ferriani, Fabrizio; Natoli, Filippo (2020). *ESG risks in times of Covid-19.* Applied Economics Letters.

Github (2017). *Fixed Effects Model.*

Gonzalez, L.; Powell, J. G.; Shi, J.; Wilson, A. (2005). *Two centuries of bull and bear market cycles.* International Review of Economics Finance.

Google (2022). *Google Trends.* https://trends.google.com/trends/explore?date=today%205-yq=ESG.

Gopinath, Gita (2020). *The Great Lockdown: Worst Economic Downturn Since the Great Depression.* International Monetary Fund.

Gress, Todd W.; Denvir, James; Shapiro, Joseph I. (2018). *Effect of Removing Outliers on Statistical Inference: Implications to Interpretation of Experimental Data in Medical Research.* Marshall Journal of Medicine.

Hansen, James; et.al. (1998). *Global Climate changes as forecast by Goddard Institute for Space Studies three-dimensional model.* NASA.

Hartzmark, S. M.; Sussman, A. B (2019). Do investors value sustainability? a natural experiment examining ranking and fund flows. *The Journal of Finance*.

ICI (2022). Investment company fact book: A Review of Trends and Activities in the Investment Company Industry.

Investopedia (2021a). Open-Ended Fund: Definition, Example, Pros and Cons.

Investopedia (2021b). Survivorship Bias.

Investopedia (2022). MSCI ESG Ratings Definition, Methodology, Example.

Jasinski, Nicholas (2020). *The SP 500 Is Trading Near an All-Time High. It's the Fastest Recovery Ever.* Barron´s.

Jesse, L. (2022). No, esg is not the same as sustainability. *KMPG*.

Kell, G (2018). The remarkable rise of esg. *Forbes*.

Kellog (u.d.). Fixed effects models. *Kellog School Of Management*.

Li, Yun (2020). *This was the fastest 30% sell-off ever, exceeding the pace of declines during the Great Depression*. cnbc.

MFS (2022). Handle with care: The role of third-party esg ratings providers in the fund selection process. *https://www.mfs.com/en-us/investment-professional/insights/sustainable-investing/the-role-of-third-party-esg-ratings-providers-in-the-fund-selection-process.html*.

Morningstar (2016). Oldest Share Class Methodology. *https://morningstardirect.morningstar.com/clientcomm/Oldest_Share_Class_Methodology.pdf*.

Morningstar (2021a). Esg investing map: How countries' stock markets stack up on esg practices and carbon risk. *https://www.morningstar.hk/hk/news/211488/map-how-countries-stock-markets-stack-up-on-esg-practices-and-carbon-risk.aspx*.

Morningstar (2021b). Morningstar sustainability rating for funds methodology. *https://www.morningstar.com/content/dam/marketing/shared/research/methodology/744156_Morningstar_Sustainability_Rating_for_Funds_Methodology.pdf*.

Morningstar (2022). Mutal funds: Data definitions: Returns. *https://awgmain.morningstar.com/webhelp/glossary_definitions/indexes/mftotalreturns.htm*.

NIST (2012). What are outliers in the data? *https://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm*.

Pastor, Ľuboš; Vorsatz, M. Blair (2020). Mutual fund performance and flows during the covid-19 crisis. *The Review of Asset Pricing Studies*.

PRI (2021). *Annual Report*. United Nations Principles for Responsible Investment.

Pustejovsky, James E.; Tipton, Elizabeth (2017). Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models. *Journal of Business Economic Statistics*.

Saeed, A., Aslam, M., Altaf, S., Amanullah, M. (2018). Improved inference for the panel data model with unknown unit-specific heteroscedasticity: A monte carlo evidence. *Cogent Mathematics Statistics*.

Sandford, Alasdair (2020). Coronavirus: Half of humanity now on lockdown as 90 countries call for confinement. *euronews*.

SEC (2022). Open-end fund liquidity risk management programs and swing pricing. *The Review of Asset Pricing Studies*.

Singer, J. D., Willett, J. B., Willett, J. B. (2003). Applied longitudinal data analysis: Modeling change and event occurrence. *Oxford university press*.

Sirri, E. R.; Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Finance*.

Taylor, Luke (2020). Why esg investors are happy to settle for lower returns. *Knowledge at Wharton*.

The Investment Association (2020). Fund statistics report december 2020.

The Investment Association (2021). Fund statistics report december 2021.

VFF (2021). Fondsundersøkelsen 2021. *Opinion*.

Appendix

A1 Figures and Tables

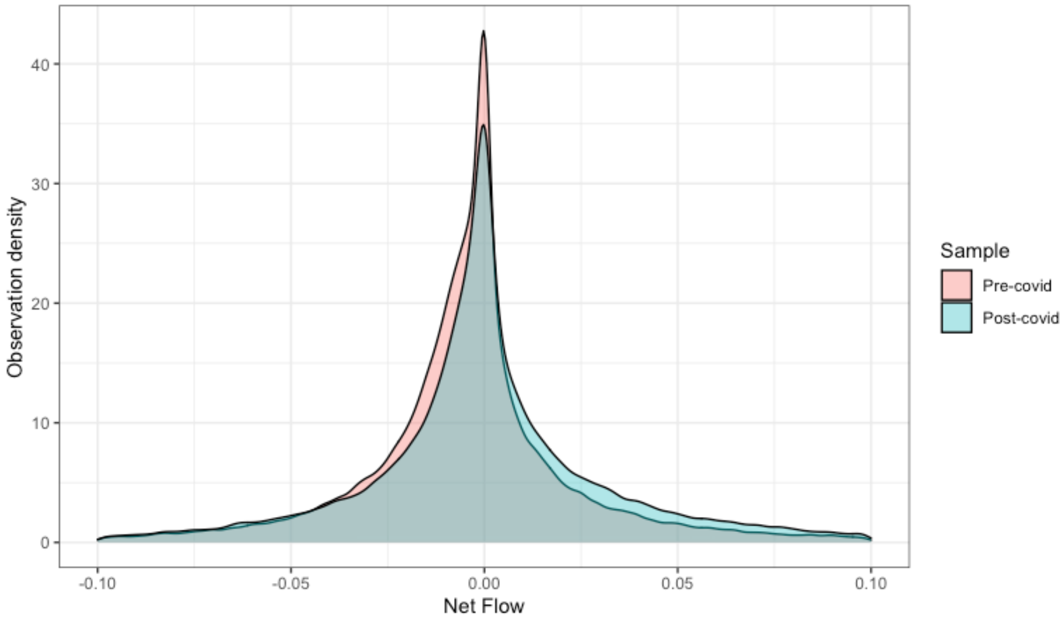


Figure A1.1: Fund Flow Density
Displays the distribution of observation density for net flow pre- and post-covid

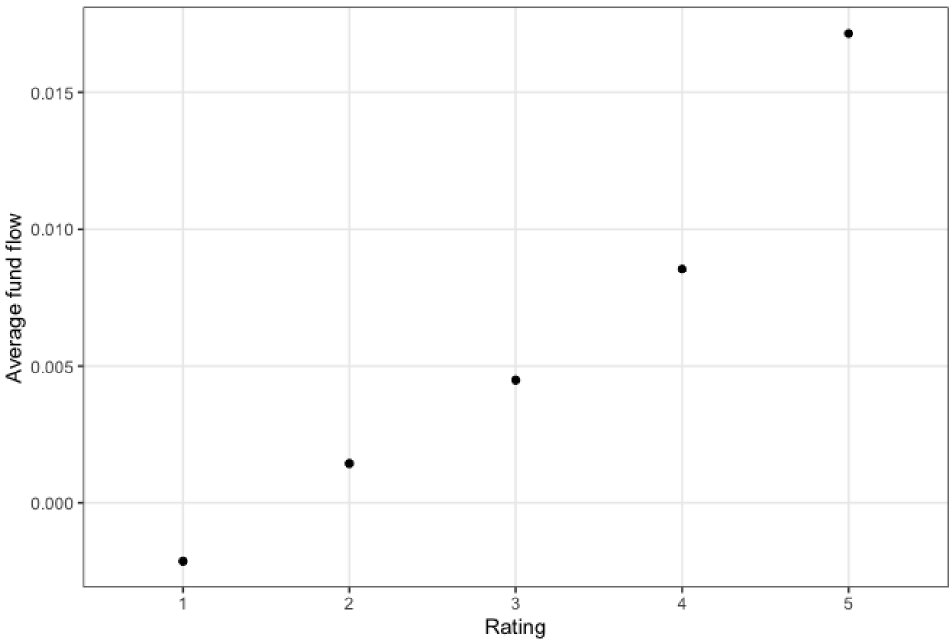


Figure A1.2: Average fund flow for each rating
Displays the average fund flow for the entire estimation period for each rating numerically. From rating 1 to 5, the average fund flow in percent is respectively -0.21, 0.14, 0.45, 0.85 and 1.7.

	Dependent variable:		
	Net Flow		
	(1)	(2)	(3)
Below average	0.002	0.007^{*}	0.004^{**}
	(0.002)	(0.004)	(0.002)
Average	0.002	0.006	0.004^{***}
	(0.002)	(0.005)	(0.002)
Above average	0.003	0.007	0.005^{***}
	(0.003)	(0.006)	(0.002)
High	0.004	0.012^{*}	0.005^{**}
	(0.003)	(0.006)	(0.002)
Observations	91,724	85,788	177,512
R^2	0.206	0.237	0.160
Adjusted R^2	0.130	0.157	0.116
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table A1.1: Fixed effects model clustered
Displays the fixed effects regressions similar to 6.2. This table is clustered on fund level.

When clustering the model, we notice a reduction of significance level for High-rated funds in the post-covid phase. The reason for the reduction is that robust standard errors are generally larger than for non-robust standard errors as it collects observations in clusters, in this case each fund, to account for heteroskedasticity. The results are still significant at the 0.1 significance level, but it is worth mentioning that clustering affects our results' significance.

Country	Code	High rated	Post-covid	DiD Coef	P-Value	P-value(High)
France	FR	0.010	-0.006	0.019	0.013	0.008
Austria	AT	0.020	0.004	0.007	0.589	0.003
Belgium	BE	0.033	-0.001	0.018	0.179	0.000
Ireland	IE	0.031	-0.005	0.015	0.094	0.000
Luxembourg	LU	0.021	0.004	0.006	0.204	0.000
Lichtenstein	LI	-0.002	0.023	-0.029	0.012	0.700
Italy	IT	0.011	0.005	0.060	0.000	0.016
Spain	ES	0.016	-0.001	0.010	0.234	0.000
Netherlands	NL	0.023	-0.016	0.030	0.327	0.132
Portugal	PT	0.005	-0.005	0.007	0.765	0.601
Sweden	SE	0.006	0.015	-0.001	0.973	0.501
Finland	FI	0.018	0.014	-0.003	0.851	0.017
Denmark	DK	0.010	0.008	0.012	0.536	0.273
Norway	NO	0.011	0.024	0.006	0.715	0.121
Germany	DE	0.005	0.011	-0.010	0.237	0.233
Great Brittan	GB	0.021	0.004	0.018	0.005	0.000
Switzerland	CH	0.018	-0.013	0.005	0.727	0.011

Table A1.2: Difference-in-difference regression table for each country
The table displays data from the difference-in-difference regressions for each country. Country is the respective country. Code is the country code for each country. High rated is the coefficient for funds that are High rated. Post-covid is the coefficient for the post-covid dummy variable. DiD Coef is the difference-in-difference coefficient. P-value is the p-value for the difference-in-difference estimators. P-value(High) is the p-value of the High rated coefficients.

A2 Data Processing

Following the data collection in Section 4.1, processing of the data is required to create data frames ready for analysis. The data from Morningstar Direct Desktop was downloaded as wide format excel-files. In line with Singer et al. (2003), this is not the preferred format for examining changes over time. The main advantage of longitudinal data is the ability to store panel data, or cross-sectional data that also captures changes over time. Since the purpose of this research is to examine changes in flows over time within mutual funds, the preferred format is longitudinal.

As the fund flow variable was not available in Morningstar Direct, it had to be calculated manually as stated in 4.4.1. Using formula 4.1 from the same section, we were able to produce fund flows. A problem when calculating the fund size percentage change data was that the first month after a fund's creation would register as an increase from zero to the first month's size. This resulted in an increase of several thousand- or million percent.

Further contaminating the fund flow calculation and leading to extreme outliers in the dataset. To correct this, an algorithm was developed to only calculate the change in fund flow if there was data (0 or not blank) for rating in period t and fund size in both period t and (t-1), if not then no value was produced. When the calculation was made for all funds, we converted the sheets to CSV files and exported them into R studios.

Panel data

After collecting all of our data, R-studios was used to transform our data into panel data, where we are able to track all the funds over the full timeframe. The code was used to create panel data for Fund Flow, Return and ESG rating across time for every fund. After merging all the data, our raw data frame was in place. This implies that we can group the data into different subsets within our dataset, being able to examine and categorise estimation periods based on all sustainability-ratings and countries.

Starting with the raw panel data, the frame was trimmed down in line with the start - and end date for the full estimation period. Then we removed index funds and unnecessary observations where either the rating or return was empty. After Winsorizing, we trimmed the panel data to fit the full sample period, which resulted in observations between August 2018 and January 2019, and March 2020 being removed. Lastly, the full timeframe was divided into two estimation periods, pre-covid and post-covid.

	FUNDS	*OBSERVATIONS*
ALL OPEN-END MUTUAL FUNDS RAW DATA	25 237	2 321 804
FILTERING:		
INDEX = “NO”	23 949	2 203 308
REMOVE N/A´S FOR RATING	10 932	371 477
REMOVE N/A´S FOR FUND FLOW	10 079	315 883
WINSORIZING	10 079	315 883
FULL TIMEFRAME	8 920	179 160
PRE-COVID	8 074	92 642
POST-COVID	8 214	86 518

Figure A2.1: Data cleaning process
Displays the changes to the dataset as cleaning measures are included.

Outliers

When analysing the converted longitudinal dataset, there were still outliers left. According to NIST (2012) an outlier is an observation that lies an abnormal distance from other values in a sample. Including outliers could introduce a higher risk of type 1 errors (Gress, 2018), weakening the chance of achieving satisfying results from the analysis. Whether the outliers present in the dataset were naturally occurring, or measurement errors, was carefully investigated. The extreme values were within fund flow due to some funds being registered as “active” their first month, with fund size near, but not quite zero. Which resulted in some extraordinary growth numbers. Omitting these values is necessary as they are problematic outliers, and not true outliers that in themselves add valuable information. These outliers would have a negative impact on the credibility of the results.