BI

WISEflow
Europe/Oslo(CEST)
03 Jul 2023

Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	09-01-2023 09:00 CET	**Termin:**	202310
Sluttdato:	03-07-2023 12:00 CEST	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202310||11184||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Fredrick Andreas Holen og Arian Odden Karlsen

Informasjon fra deltaker

Tittel *:	ESG rankings, Investor Preferences, and Net Flows: Evidence from US Mutual Equity Funds in Market Downturns
Navn på veileder *:	Bruno Gerard

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	348
Andre medlemmer i gruppen:	

- ESG rankings, Investor Preferences, and Net Flows: Evidence from US Mutual Equity Funds in Market Downturns -

Date of submission:
03.07.2023

Supervisor:
Bruno Gerard

Campus:
BI Norwegian Business School, Oslo

Examination code and name:
GRA 19703 Master Thesis

Programme:
Master of Science in Business, Major in Finance

Abstract

This study examines the relationship between fund sustainability and net flows of open-ended US mutual equity funds, and investor preferences in market downturns. We find evidence of a pattern using panel data fixed effects models and incorporating ESG scores from multiple providers. During market downturns, investors tend to withdraw from funds with weak sustainability performance and gravitate towards those with strong sustainability performance. This preference for sustainability holds for both the COVID-19 pandemic and the Russia-Ukraine War. Overall, the study provides valuable insights into the relationship between fund sustainability, investor behavior, and market conditions, highlighting the significance of sustainability as a factor influencing investment decisions in times of uncertainty.

Acknowledgments

We would like to thank our supervisor Bruno Gerard for his support, guidance, and insights during the whole process. Additionally, we would like to thank our families and friends for their unconditional support.

Table of Contents

List of Tables

List of Appendix

1. Introduction and Motivation

In this paper, we focus on examining the effect of sustainability on fund flows for US open-ended equity mutual funds, with a particular emphasis on understanding its influence during market downturns. Sustainable investing has experienced remarkable momentum and popularity in recent years, becoming a prominent force in the global financial landscape. Bloomberg (2022) reports that globally, money invested in sustainable funds witnessed an astonishing growth rate of approximately 50% from 2020 to 2021. This surge in interest and investment led to a substantial increase in inflows to sustainable funds. McKinsey (2022) notes that from $5 billion in 2018, inflows to sustainable funds skyrocketed to over $50 billion in 2020 and expanded to nearly $70 billion in 2021. These figures demonstrate a significant influx of capital into sustainable investment vehicles, reflecting the rising demand for sustainable investment opportunities.

The COVID-19 pandemic and the Russia-Ukraine War have significantly impacted global markets and economies. The COVID-19 pandemic resulted in widespread restrictions and a stock market crash in February. The Russia-Ukraine War has contributed to global economic insecurity, where the conflict resulted in a spike in energy prices, exacerbating an already challenging inflationary situation. Given the significance of these market downturns, we find it interesting to examine their impact on fund flows concerning ESG considerations.

Numerous studies have highlighted the significance of sustainability on fund performance and investor preferences (e.g., Pastor and Vorsatz (2020), Ammann et al. (2018), Hartzmark and Sussman (2019), Renneboog et al. (2008), Bollen (2007)). Furthermore, research conducted on crises has indicated an investor preference for sustainable funds, suggesting an observed "flight to safety" effect characterized by increased demand for sustainability (see Pastor and Vortzas (2020), Ferriani and Natoli (2020), Parida and Wang (2008)). We contribute to the ongoing research by investigating the fund flows of high-ranking ESG funds and low-ranking ESG funds compared to the average fund. In contrast to many other studies that predominantly emphasize returns as a measure of performance, our research takes a distinctive approach by focusing primarily on fund flows.

Additionally, we incorporate the impact of the Russia-Ukraine War as a market shock, providing a unique context for our analysis.

Our thesis will propose to answer the following questions:

1. *Does fund ESG rank affect the net flows of US mutual equity fund?*
2. *Does high-ranking ESG funds and low-ranking ESG funds experience differences in net flows compared to average ESG funds in market downturns?*
3. *Does high-ranking ESG funds and low-ranking ESG funds experience differences in net flows compared to average ESG funds in Covid-19 and Russia-Ukraine conflict?*

To address these questions, using a sample covering the period from 2012 to 2022, we calculate the fund-level ESG score by matching fund holdings with individual stock ESG scores and followingly conduct multiple fixed effects regressions controlling for relevant factors influencing fund flows. This allows us to study the influence of fund ESG rank on the net flows of US mutual equity funds.

Our study reveals that open-ended US equity mutual funds do not exhibit significant differences in net flows over the entire period or during normal market conditions. However, we observe evidence that investors place a higher value on sustainability during market downturns and tend to avoid funds with poor sustainability performance. This trend holds for the periods of market downturn we analyzed, namely the Covid-19 pandemic and the Russia-Ukraine War.

2. Literature review

This section is divided into four distinct parts: the determinants of fund flows, ESG disagreement and classification, sustainability, investor behavior and mutual funds, and sustainability in crises.

2.1 Determinants of Fund Flows

In relation to the study of mutual funds, it is vital to understand the determinants of fund flows. In this section, we review the evidence about the main drivers of fund flows that we would need to control for in our empirical investigation to isolate the effect of sustainability on flows. Warther (1995) finds that fund inflows and returns are positively related. This implies that aggregate security returns are strongly correlated with unexpected concurrent net flows into mutual funds and are uncorrelated with concurrent expected net flows. Furthermore, the article illustrates a correlation between the fund flow and the return of the securities held by the fund, i.e., the flows into stock funds are correlated with the underlying stock return.

In another study examining the fund flow determinants in mutual funds and pension funds, Del Guercio and Tkac (2000) argue that the mutual fund flow performance relation is highly convex. This implies that mutual fund investors disproportionately flock to good performers but do not punish poor performers by withdrawing assets. In comparison, pension funds punish poor performance by withdrawing assets. The authors argue that mutual fund investors may be a relatively unsophisticated client base, as they flock disproportionately to recent winners and use raw return performance. Further, the study finds that Mutual fund manager flow is significantly positively correlated with Jensen's alpha, which is interesting considering the client base to be unsophisticated. This is explained mainly by the clients' access to summary performances, such as Morningstar Star ratings, which are highly correlated with this Jensen's Alpha. Finally, the study shows that significant mutual funds attract flow approximately in proportion to their asset size.

In a similar study, Del Guercio and Tkac (2008) find evidence that Morningstar Star rating affects investor flows independently of the influence of other standard measures of fund performance. They find that the discrete change in the star rating itself is not the change in the underlying performance measures that drive mutual fund flows. The intuition is that the star rating is easy to access and low cost for investors, many of whom would otherwise find fund selection intimidating or overwhelming.

Sirri and Tufano (1998) argue that equity mutual fund inflows are sensitive to historical performance, although not in a linear manner. For mutual funds in the top quantile of funds in their objective category, the performance is positively correlated and results in economically and statistically significant inflows. Whereas for the lowest-performance funds, the relationship is virtually non-existent. Furthermore, the article finds evidence that a reduction in annual fees has a significant positive effect on the flows, even though low-cost funds have fewer available resources to use on the market this performance.

2.2 ESG-rating Disagreement

ESG rating varies across providers, and fund-stated objectives can differ from the actual reflection of the fund holdings. Results from Berg et al. (2019) study on six prominent ESG rating providers document rating divergence. The study revealed an average correlation of 54% among the various providers, ranging from 38% to 71%. In a similar study, Gibson et al. (2021) collected a unique sample with a comprehensive data coverage of ESG ratings. They find the average correlation between the ESG ratings of the providers to be 45%. Additionally, the researchers examine whether the level of disagreement varies based on observable financial and accounting characteristics of individual firms. Their findings suggest that disagreement is higher among the largest firms in the S&P 500, potentially due to the complexity associated with such companies.

2.3 Sustainability, Investor Behavior, and Mutual Funds

Several studies investigate the effects of sustainability and mutual fund flows. In line with Del Gurierco and Tkac's (2008) findings on the effect of Morningstar rating, a study by Hartzmark and Sussman (2019) involved a natural experiment on the introduction of Morningstar's ESG globe rating in 2016. The paper found that mutual fund investors collectively treat sustainability as a positive fund attribute, allocating more money to funds ranked five globes (highly sustainable) and less money to funds ranked one globe (less sustainable). Although this event does not universally explain the investor preference for sustainability, the experimental evidence suggests that investors view sustainability as positively predicting future performance. A similar study conducted by Ammann et al. (2018) finds evidence supporting retail investors shifting their money towards high-rated funds, from low-rated funds, in the event of the shock related to the release of Morningstar's sustainability rating.

Despite Hartzmark and Sussman's findings, they find no evidence of high-sustainability funds outperforming low-sustainability funds. Also, the study illustrates that the investor cares more about this simple rating than all underlying information. Such information was present before but ignored, making the sustainability aspect more accessible in their decisions.

In relation to the volatility of fund flows, Bollen (2007) provides evidence of lower monthly volatility of investor cash flows in socially responsible funds compared to conventional funds. Hence, socially responsible investment (SRI) fund flows exhibit a higher (lower) level of sensitivity to past positive (negative) returns in comparison to conventional funds. This difference is shown to be robust over time and persistent as funds age. However, the research of Renneboog et al. (2011) and Benson and Humphrey (2008) reveal that US SRI fund flows are less sensitive to performance when compared to their matched conventional funds. Consequentially, this addresses the issue stated by Markowitz (1952). The author suggests that social screens might serve as filters for management quality and hence generate superior risk-adjusted returns. In the context of investor behavior, Bollen argues that investors may have a multi-attribute utility function that is not exclusively based on the standard risk-reward optimization but also incorporates a set of personal and societal values.

Renneboog et al. (2008) find that volatility in socially responsible funds is lower than conventional funds flow volatility. If the utility function of the investor is based on SR values, one should expect "(i) further SRI growth even if the risk-adjusted SRI returns are lower than those of conventional investments, and (ii) less sensitive SRI money-flows to past performance." The argument that investors care about non-financial attributes of their investments (like sustainability) is consistent with previous literature. In contrast to Nofsinger and Varma's (2014) conclusion, Renneboog finds no difference in the alphas for the inflow and outflow portfolios of socially responsible mutual funds and conventional mutual funds domiciled in the United States.

As previously mentioned, Del Guercio and Tkac (2000) argue that the relationship between mutual fund flows and performance is highly convex, indicating that investors tend to flock to good-performing funds but do not penalize poor performers by withdrawing their investments. In the context of socially responsible investment (SRI) funds, this implies that investors are more likely to keep their money invested in SRI funds, even if they underperform compared to conventional funds (Nofsinger and Varma, 2014). This behavior can be attributed to the perception that it is better to experience smaller losses during market downturns than to chase larger gains during market upturns. Therefore, investors may opt for a portfolio with asymmetric performance, as the utility gained from performing better in falling markets outweighs the utility lost from underperforming in rising markets.

2.4 Sustainability in Downturns

In our study, it is imperative to understand sustainability in the context of crisis. Nofsinger and Varma (2014) investigate the performance of socially responsible funds and market crises. The study shows that socially responsible mutual funds outperform matched conventional mutual funds during the two crisis periods; the technology bubble burst and the global financial crisis. Although socially responsible mutual funds generate a reduction of downside risk, such funds underperform during non-crisis periods. The analysis shows that the outperformance in crisis periods is driven by the mutual funds that focus on shareholder advocacy and ESG issues. Conclusively, in market crisis periods, positive socially responsible attributes of businesses make them less risky. This supports Oikonomou et al. (2012) theory that socially responsible behavior is weakly negatively related to systematic risk, while irresponsible behavior is strongly positively related to systematic risk. The authors conclude that the difference in performance is a result of the socially responsible characteristics of the underlying stocks and not necessarily the portfolio management or stock-picking ability of the fund.

Parida and Wang (2018) examine Corporate Social Responsible (CSR) mutual fund flows prior to, during, and after the global financial crisis in the US. The study finds that top CSR funds attract less annual investments (about 5%) compared to other funds during the period from 2003 to 2012, but in the financial crisis period, top CSR funds receive more investments (about 9%) compared to the pre-crisis period. Whereas bottom CSR funds attract more investments (about 6%) compared to other funds during the period from 2003 to 2012, and in the financial crisis periods, bottom CSR funds receive fewer investments (about 10%) compared to the pre-crisis period. However, this increases in investments in top CSR funds does not sustain after the crisis period. Their findings imply that in the presence of financial stress, investors perceive top CSR funds as relatively high quality or safety, as more investments flow into these funds in the financial crisis. This finding is consistent with the "flight to quality" phenomenon.

Pastor and Vorsatz (2020) relate fund performance and sustainability. Their findings suggest investors show a preference for funds that implement exclusion criteria and have high sustainability ratings, particularly environmental funds. Pastor and Vorsatz find that during the COVID-19 crisis, funds that receive high sustainability ratings and those with high star ratings demonstrate strong performance. The researchers further oppose the classical perception of the environmental quality feature being a "luxury good" (Baumol and Oates, 1979), where their results exhibit that investors now perceive sustainability as an essential requirement rather than a discretionary luxury.

Additionally, Ferriani and Natoli (2020) shed light on the impact of ESG risk during the COVID-19 crisis, identifying a "flight-to-safety" effect towards low-ESG risk funds. Their study employing Morningstar's newly introduced ESG risk shows that investors have notably favored low-ESG risk funds while discarding high-risk ones since the financial markets experienced a crash in late February 2020.[1] This indicates that the COVID-19 crisis has heightened the significance of ESG risk considerations among investors.

[1] Morningstar introduced their "ESG risk indicators" at the end of 2019.

Dottling and Kim (2022), investigating the impact of COVID-19 on SRI mutual fund flow, provide evidence that funds with higher sustainability ratings tend to encounter more pronounced reductions in retail fund flows, using COVID-19 as an economic shock. They further fill the uncovered gap in the literature regarding the sensitivity of SRI demand among retail investors in relation to changes in economic conditions. They find fragility in demand in SRI for retail investors.

3. Background and hypothesis

In this section, we discuss relevant subjects to substantiate why our research question is interesting to investigate.

3.1 Environmental, Social, and Governance

In the space of finance, sustainability is often contemplated as Environmental, Social, and Governance, in short, ESG. The ESG factors play a crucial role in evaluating a portfolio's sustainability across various aspects and assist investors in aligning their ESG criteria with the principles of sustainable development when assessing potential investment opportunities (Boffo & Patalano, 2020). ESG investing involves the screening of investments based on corporate policies and aims to incentivize responsible behavior among companies, mutual funds, and other investment products (Investopedia, 2023). ESG investing can be interpreted as market participants collectively pursuing a shared objective known as "green investing." In accordance with De Spiegeleer et al. (2020), we observe stakeholders considering the ESG dimensions at their core and intent to enhance companies or portfolios within the three dimensions. Similarly, in relation to mutual fund investors, retail investors value sustainability in fund investing and treat sustainability as a positive fund attribute (see, e.g., Hartzmark and Sussman (2019), Ammann et al. (2018); Pastor and Vortsaz (2020), Renneboog et al. (2008), Nofsinger and Varma (2014), Bollen (2007))[2].

On the other side, the empirical evidence presents conflicting results regarding the impact of sustainability yield in the context of fund performance. Hartzmark and Sussman (2019), Humphrey et al. (2016), and Renneboog et al. (2008) find no evidence of a significant relationship between fund performance and sustainable

[2] We mention «sustainability», although the articles utilize different measures such as *socially responsible investments and funds*.

investing. Nofsinger and Varma (2014) and Jones et al. (2008) suggest that sustainable funds generally underperform compared to conventional funds.

The existing body of ESG research on equity mutual funds predominantly focuses on fund return as a performance metric, while the study of cash flows into and out of funds remains relatively limited. In light of this gap, our aim is to contribute to this area of research by examining the net flows of funds and exploring the impact of ESG factors on fund performance. By investigating the relationship between ESG and fund flows, we seek to enhance the understanding of how ESG considerations influence investor behavior and the overall dynamics of the funds' financial performance. Additionally, we aim to shed light on the contradictory empirical evidence regarding the performance of sustainable funds with respect to non-sustainable and average funds.

During the past decade, investor preferences have been notably influenced by ESG factors. We find it intriguing to investigate the ESG dimension within mutual funds to gain valuable insights into the implications of sustainability in the mutual equity fund market and its influence on retail fund investments.

3.2 COVID-19

On January 30th, 2020, the World Health Organization (WHO, 2020) officially declared the COVID-19 epidemic as a global public health emergency following reports of over 7,000 cases worldwide. Subsequently, the number of cases began to increase, leading to the declaration of COVID-19 as a pandemic on March 11th (WHO, 2020). By the end of March, nearly half of the world's population was under restrictions (Sky News, 2020).

The crisis disturbed the financial markets and economy with unprecedented speed, causing a stock market crash in the middle of February (Baker et al., 2020). Consequently, the financial markets transitioned into a bear market, defined as a decline of at least 20 percent from the previous peak index level (Gonzalez et al., 2005).

The significance of sustainability became prominent during the market crash, leading to the publication of articles focusing on ESG. As proposed by Pastor and

Vorsatz (2020), investors show a preference for sustainable funds during the COVID-19 crisis, and Ferriani and Natoli (2020), the investor demand for low-ESG risk funds increased while the high-ESG risk is discarded, implying a "flight to safety" effect. On the contrary, Dottling and Kim (2022) show that funds with higher sustainability ratings experience more pronounced reductions in retail fund flows. The importance of ESG during the initial phase of the COVID-19 pandemic was highlighted as early as April 2020, when ESG investments exhibited resilience and outperformed other holdings (Morningstar, 2020). In line with these articles, our study focuses on examining the comprehensive impact of the COVID-19 pandemic on fund flows specifically related to ESG.

3.3 Russia-Ukraine War

The prolonged and intense conflict in Ukraine has inflicted devastation upon the country, deepened the divide between Western nations and Russia, and contributed to global economic insecurity (New York Times, n.d.). The onset of the Russian invasion of Ukraine had an instantaneous and profound impact on global markets, leading to significant losses in stock market indices (Izzeldin et al., 2023). The escalation of the conflict resulted in a spike in energy prices, directly affecting consumers and industries with high energy-dependent costs, especially for countries heavily reliant on energy imports from Russia (MSCI, 2023). This surge in energy prices exacerbated an already challenging inflationary situation, which was partly attributed to the monetary and expansionary fiscal policies implemented during the peak of the COVID-19 crisis. As a result, the global economy experienced a substantial shock, with energy and food markets experiencing the full impact of the disruption, leading to severe shortages in supplies and extraordinary increases in prices (ECB, 2023).

Despite the limited existing research on the impact of the Russia-Ukraine conflict on mutual fund sustainability and fund flows, we seek to scope current body knowledge and contribute to this area of study. Our objective is to investigate the effects of this conflict on sustainable mutual fund flows, thereby enhancing our understanding of the relationship between fund flows and sustainability during times of crisis.

3.4 The US Market

The US market is characterized by its significant size and inhabits various similarities within the country. The US markets share characteristics, such as market structure and regulatory environment, which further enhance the similarities among equity funds. Studying equity funds in the US provides valuable insights into the dynamics of fund flow. In our study, we specifically focus on equity funds as they exhibit greater volatility and cross-sectional variation (Warther, 1995), making them an ideal context for examining the dynamics of fund flow. Concentrating on equity funds allows us to explore the patterns and behaviors associated with the flow of funds in a more comprehensive and nuanced manner.

3.5 Hypothesis

Based on our discussion of literature and background, we formulate the following hypotheses to guide our study:

Hypothesis 1:

H_0: High and Low ranked funds do not receive a difference in Net flow compared to average ranked ESG funds

H_1: High and Low ranked funds do receive a difference in Net flow compared to average ranked ESG funds

Hypothesis 1 is created in line with previous literature finding evidence that ESG factors positively influence the net flows of mutual funds. Based on the increasing popularity of ESG investing and the emphasis on sustainability by investors, we expect that mutual funds with stronger ESG performance will attract higher net flows compared to those with weaker ESG performance. This hypothesis aligns with the notion that investors value sustainability as a positive fund attribute and are more inclined to invest in funds that align with their ESG criteria (Hartzmark and Sussman (2019); Bollen (2007); Renneboog et al. (2008)).

Hypothesis 2:

H_0: High and Low ranked funds do not receive a difference in Net flow compared to average ranked ESG funds in market downturns

H_1*: High and Low ranked funds do receive a difference in Net flow compared to average ranked ESG funds in market downturns*

Hypothesis 2 is based on the arguments presented by Nofsinger and Varma (2014), Pastor and Vorsatz (2020), and Parida and Wang (2018), which suggest that High ESG funds exhibit superior performance compared to average funds during market downturns.

Hypothesis 3:

H_0*: High and Low ranked funds do not receive a difference in Net flow compared to average ranked ESG funds in COVID-19 and the Russia-Ukraine Conflict.*

H_1*: High and Low ranked funds do receive a difference in Net flow compared to average ranked ESG funds in COVID-19 and the Russia-Ukraine Conflict.*

Hypothesis 3 examines the individual effects of the COVID-19 crisis and the Russia-Ukraine Conflict on the net flows of funds characterized by stronger ESG performance and weaker ESG performance. The hypothesis is built upon the findings of Pastor and Vorsatz (2020) and Ferriani and Natoli (2020), which provide evidence of a positive association between sustainability and fund flows during the COVID-19 crisis. In contrast, the study by Dottling and Kim (2022) indicates that funds with high sustainability rankings experience significant declines in fund flows during the COVID-19 pandemic.

4. Empirical Methodology

In this section, we develop a robust framework to answer our testable hypotheses. We discuss potential approaches to investigate fund flows, carefully selecting the most suitable approach.

4.1 Model Specifications

In this study, our objective is to examine the impact of sustainability on fund flows for open-ended US equity mutual funds. To achieve this, we will specify a measure of ESG performance and define the periods of interest. We create rankings based on the ESG score from different providers to capture the effect of ESG performance on fund flows. This enables us to determine whether there are differences in fund flows between funds with strong and weak ESG performance. To account for the potential influence of other variables associated with fund flows, we introduce control variables suggested by previous literature. This will allow us to isolate the effect of ESG performance on fund flows while controlling for other factors.

We include monthly returns to capture the relationship between past performance and fund flows, as highlighted by Warther (1995), Del Guercio and Tkac (2000), and Sirri & Tufano (1998). Size is included as Del Guercio and Tkac (2000) found evidence of its impact on fund flows. Morningstar ratings serve as a control variable to reflect the influence of summary performance measures on fund flows, as shown by Del Guercio and Tkac (2008). We also incorporate annual expenses to account for the positive effect of reducing fees on fund flows, as suggested by Sirri & Tufano (1998). Additionally, risk is considered through the inclusion of standard deviation. The logarithm of fund age is included to capture the potential effect of fund experience.

In some cases, it is expected that the effect of the dependent variable in relation to one explanatory variable may vary depending on the level or magnitude of another explanatory variable. This phenomenon is known as an interactive effect or interaction effect (Wooldridge, 2021). In relation to studies on fund performance and ESG (see, e.g., Nofsinger and Varma (2014), Pastor and Vortsaz (2020), Renneboog et al. (2008), and Barko, Cremers and Renneboog (2021)), implies the possibility of an interaction effect between fund's ESG score and return. Therefore, an interaction term between ESG rankings and monthly return is introduced to capture potential joint effects on fund flows. Likewise, studies related to fund expense and ESG (see, e.g., Kempf and Osthoff (2008)) also suggest the presence of an interaction effect. Resulting in an interaction term between ESG measures and expenses also included.

To address our hypotheses, we will conduct multiple fixed effects regressions to examine the relationship between ESG and fund flows. Initially, we would perform a regression analysis over the entire sample period. This will allow us to establish a baseline understanding of the relationship between ESG and fund flows within our sample. However, due to the potential high correlation between ESG rankings from providers, we will conduct separate regressions for each of the agencies to explore the impact of ESG. The baseline regression equations for this model are as follows:

$$(1)\quad Net\,Flow_{i,t} = \beta_1 HighESG_{i,t-1} + \beta_2 LowESG_{i,t-1} + \theta I_{i,t-1} + \gamma X_{i,t} + \mu_i + \eta_t + \varepsilon_{i,t}$$

Where $Net\,Flow_{i,t}$ denotes percentage inflow or outflow into fund *i* at period *t*. $HighESG_{i,t-1}$ and $LowESG_{i,t-1}$ represents dummy variables indicating if fund *i* has a high ESG-score rank (top 10%) or a low ESG-score rank (bottom 10%), in the previous period. . $I_{i,t-1}$ is a vector containing the interaction terms between the ranked funds and return/expense, for fund *i* in the previous period. $X_{i,t}$ is a vector comprising of the fund control variables; $Return_{i,t-1}$, $Age_{i,t}$, $Size_{i,t-1}$, $Expense_{i,t-1}$, $MSR_{i,t-1}$, and $Std\,Dev_{i,t}$. μ_i is a dummy variable for fund *i*, representing the fund fixed effects which allows the control for fund-specific factors that are time-invariant. We also control for time-specific factors, η_t, which is a dummy variable representing the time fixed effects. While $\varepsilon_{i,t}$ is the error term.

To further analyze the effect of ESG rankings on fund flows, we separate the sample into two subsamples: crisis periods and normal periods. This allows us to examine the relationship between net flows and sustainability rankings during different time intervals and differentiate the effect of ESG between normal periods and market downturns. We conduct regressions equivalent to equation (1) for each subsample.

For the first two sections of our analysis, the coefficients of interest are β_1 and β_2, which serve a crucial role in describing the estimated effects of a High ESG and a Low ESG rank on net flows, relative to those of the average fund. Positive and statistically significant coefficients suggest that funds within the rank have a higher inflow compared to funds with an average ESG rank. Conversely, a negative and statistically significant coefficient indicates that funds within that category experience higher outflows relative to funds with an average ESG rank. In the sense of our hypothesis, the significance for β_1 or β_2 will lead to the hypothesis being rejected, indicating one of the rankings influencing fund flow.

Furthermore, we perform an additional regression analysis incorporating dummy variables for both the Covid-19 pandemic and the Russia-Ukraine conflict. This approach follows a similar methodology employed by Dottling & Kim (2022), where they utilize a dummy variable to capture the impact of Covid-19. In our regression, we extend this by including the variable "Conflict" to account for the Russia-Ukraine conflict. By incorporating these dummy variables, we will be able to analyze the effect of ESG ranks on net flows during individual crises.

$$
\begin{aligned}
(2)\ Net\ Flow_{i,t} = {} & \beta_1 HighESG_{i,t-1} \times Covid_t + \beta_2 LowESG_{i,t-1} \times Covid_t \\
& + \beta_3 HighESG_{i,t-1} \times Conflict_t + \beta_4 LowESG_{i,t-1} \times Conflict_t \\
& + \beta_5 HighESG_{i,t-1} + \beta_6 LowESG_{i,t-1} + \theta I_{i,t-1} + \gamma X_{i,t} + \mu_i + \eta_t + \varepsilon_{i,t}
\end{aligned}
$$

Where $Covid_t$ and $Conflict_t$ serves as an indicator for the months within our defined crisis periods. Taking on the value of 1 if the date falls within their respective month, zero otherwise. We are interested in the interaction effect between the sustainability ranks and the two crisis dummies on net flow, denoted as $\beta_1,\ \beta_2, \beta_3, and\ \beta_4$. The coefficients provide estimates of the difference in net flows between the rankings and average ESG funds separated into Covid and Conflict. A positive and statistically significant coefficient suggests that funds within that rank experience higher inflows compared to funds with an average ESG rank during that period. Conversely, a negative and statistically significant coefficient indicates that funds within that rank experience higher outflows relative to funds with an average ESG rank during that period. Significance for coefficients $\beta_1,\ \beta_2, \beta_3, or\ \beta_4$, would lead to the rejection of the null hypothesis, indicating that the rankings influence fund flows in at least one of the periods.

4.2 Model Selection

This study will utilize panel data regressions to explore the relationships between ESG performance and fund flows. Our data will consist of both time-series and cross-sectional elements, and the use of panel data regressions offers the possibility of examining the relationships between the variables dynamically while accounting for other controls. For panel data, there are three main types of models to consider, Pooled-OLS, Fixed-Effects (FE) model and the Random-Effects (RE) model (Brooks, 2019). Determining the appropriate model depends on the underlying assumptions and the nature of our data.

According to Wooldridge (2010), when analyzing panel data, Pooled OLS can be used when a different sample is selected for each period. This approach essentially pools the data from all periods and treats it as a single cross-sectional dataset. In our analysis, we are working with the same funds across the periods, indicating that Pooled OLS might not be appropriate for our data. We assume that a FE model is a better approach than a RE model when analyzing the effect of ESG on fund flows, as we believe that there are unobserved individual or group-specific effects that are correlated with the independent variables. FE estimator can address both fund- and time-specific effects by capturing the unit- and time-invariant differences between the funds and months (Brooks, 2019), denoted as μ_i and η_t, where μ_i are entity- and η_t are time-specific. Unlike the other estimators, the FE estimator does not include a general constant term. Instead, the individual effects are captured by the dummy variables for each entity and time period in the regression, enabling us to capture the entity-specific and time-specific effects on the dependent variable. To further test our assumptions and decide which model is preferred, we will also conduct a restricted F-test and Hausman test (Brooks, 2019).

4.3 Robustness of result

Based on the methodologies from previous literature on fund flows, we assume our data will suffer from heteroscedasticity, and to check for this, we run a Breusch-Pagan Test, where a rejection of H_0 reveals the presence of heteroscedasticity within our sample. Additionally, we also assume autocorrelation will be a problem for the sample, and a Breusch-Godfrey Test is also conducted to check for the existence of autocorrelation. Consequently, if our assumptions show to be true, we account for these issues by the utilization of Driscoll-Kraay standard errors (HAC) in our regressions, as proposed by Driscoll and Kraay (1998)[3]. This decision will enhance the robustness of our statistical inferences, as well as improve the comparability and consistency of the results across models, thereby making our models more conservative.

Multicollinearity can complicate the interpretation of regression results when working with panel data. To mitigate this, we calculate the Variance Inflation Factor (VIF). A VIF value of 10 or above is considered problematic for statistical inference. While multicollinearity does not compromise the properties of Best Linear Unbiased Estimators (BLUE) in regression, according to Brooks (2019), high VIF values for the variables of interest, such as ESG rankings, can make it challenging to interpret their individual effect. Therefore, we strive to keep the VIF of these variables relatively low.

For the control variables, we include those that do not contribute high VIF to our ESG rankings. We do not find it problematic to have high VIF for the control variables if they demonstrate statistical significance with plausible magnitudes and appropriate signs. This approach aims to enhance the validity of our regression analysis. It aligns with established practices and prior research emphasizing the importance of these control variables in predicting net flows.

To test the robustness of our results, we plan to conduct additional regression analyses with different modifications. We will include fund family fixed effects to account for unobserved heterogeneity across fund families. Additionally, we will explore broader definitions of our rankings to assess the sensitivity of our estimates. These modifications allow us to examine the influence of our assumptions on the observed relationship and ensure the reliability of our findings.

1. Fund family fixed effects: Instead of using fund fixed effects, we use fund family fixed effects to control for unobserved heterogeneity at the fund family level. This would allow us to see if the results still hold when controlling for differences between fund families rather than differences between individual funds.

---

[3] We follow the approach of many practitioners and use bandwidth as $T^{\frac{1}{4}}$ (Greene, 2011, p. 960)

2. ESG rankings of 20%: Instead of using 10% as a threshold, we increase the Low and High to 20% limits. This would allow us to see if the results still hold when using a broader definition of High and Low ESG funds.

5. Data

This section provides a description of the data used in the study. The following sections provide a detailed explanation of the collection process, screening, and construction process of the mutual fund data and the ESG data on a mutual fund level. Additionally, we implemented a discussion of potential weaknesses in the dataset.

5.1 Data Collection Process

The primary data sources in this study are Morningstar and Eikon Refinitiv, specialized providers of fund-level and company-level data, respectively. Morningstar Direct is Morningstar's extensive database for mutual funds and offers survivorship bias-free data, covering essentially nearly all variables needed in the analysis. Refinitiv is used to obtain the ESG score of the individual companies in each of the funds' portfolios to create an artificial ESG measure we use in the analysis. This is due to the limitations of the ESG measures offered by Morningstar for fund-level assessments prior to the end of 2018.

To arrive at our final sample, we followed a series of steps. First, we select all open-ended US mutual equity funds that primarily focus on investments within the US, which is in line with the scope of our thesis. This resulted in a pool of approximately 12,000 fund-share classes available in the Morningstar Direct Database. Next, we filtered out fund share classes that did not have at least one observation for the variables throughout the entire sample period, reducing the sample to 4 429 fund share classes. This ensured that we only included share classes with complete data for analysis.

Finally, we extracted the required variables for the funds included in our final sample. These variables comprised monthly return, standard deviation, inception date, Morningstar rating, net assets, and expense ratio on the share class level. We also obtained net flow and total net assets already aggregated to the fund level. Additionally, we obtained fund portfolio holdings as they were necessary for calculating the ESG score for each fund.

5.2 Sample Construction

Our analysis is conducted at the fund level, with data aggregated using value weights. This was accomplished by dividing the net asset value of each share class by the total net asset value of the fund within the same period. In cases where a variable was not available for a specific fund share class, we used the average value of the other share classes within the same period to calculate the weighted average. It is important to note that the inception date of the fund corresponded to the oldest share class within that fund. While Size and Net flow are already at the fund level, we also address incubation bias by excluding funds with assets under management of fewer than 5 million USD and an age of less than one year.

5.3 Sample Period

The sample period for our analysis spans from January 2012 to December 2022, with a monthly frequency. Within this timeframe, we have identified two significant events: the COVID-19 pandemic and the Ukraine-Russia war. To capture the variations in fund flows between periods of stability versus periods of market downturns, we divide our sample into sub-samples representing normal and crisis periods.

The normal sub-sample covers two time periods: January 2012 to February 2020 and June 2020 to February 2022. These periods are considered relatively stable and unaffected by the major crises we are examining. The crisis sub-sample, on the other hand, encompasses two periods: March 2020 to May 2020[4], and March 2022 to May 2022[5]. These crisis sub-samples start from the first period following the occurrence of the respective events, as our data reflect fund characteristics at the beginning of each month. This ensures that the effects of any outflows or other impacts are accounted for in the subsequent periods.

[4] Initial three months after the World Health Organization's declaration of COVID-19 as a health emergency on January 30, 2020

[5] Initial three months after Russia's invasion of Ukraine

5.4 Sustainability Data

To determine the equity mutual fund sustainability score, we access the fund portfolio's underlying stock-level ESG scores. In the combination of the stock's value-weight and standardizing ESG stock ratings, we calculate the fund-level ESG score based on ratings from Refinitiv and Sustainalytics individually. With the inspiration of Gibson and Krueger (2018), we aim to compute the fund ESG score and assign the scores to High-ranking ESG funds (top 10%) and Low-ranking ESG funds (bottom 10%).

4.4.1 The data providers

The two primary data sources for ESG data in the study are the Refinitiv Eikon database and the Morningstar Sustainalytics database. Our focus will be directed toward the aggregated score of the three pillars Environmental, Social, and Governance. The two providers have developed their own ESG score and methodology, but the scores range on a scale from 0 to 100 in their assessment. Due to differences in methodologies between the providers, we evaluate the scores independently. We collect ESG data on a company level for publicly listed firms from 2012-2022, listed on stock exchanges in Asia, Europe, Africa, the Americas, and Oceania.

Followingly we also collected companies missing from these listings, which are present in the fund holdings. Implementing screening procedures resulted in 11 959 fund-year observations. To better match the fund holdings from Morningstar, we collected a substantial number of firms, although Refinitiv and Sustainalytics were not able to provide the necessary data for all firms in our sample.

4.4.1.1 Sustainalytics ESG Score

Sustainalytics' ESG Risk Ratings offer a sophisticated and comprehensive approach to measuring a company's exposure to industry-specific material ESG risks and its ability to manage those risks. This multi-dimensional methodology combines considerations of management practices and exposure levels to provide an absolute assessment of ESG risk. The ESG risk ratings provide comprehensive coverage, including more than 16 300 ESG analyst-based ESG risk ratings. With a transparent methodology, Sustainalytics provides access to multiple levels of data and qualitative insights. Followingly, Sustainalytics follows an annual update cycle to ensure that the ESG Risk Ratings remain current and relevant.

Sustainalytics employs a proprietary model to assess a company's exposure to ESG risk across 138 sub-industry classifications. They consider the impact of 20 material ESG issues (MEIs), selecting up to 10 MEIs per sub-industry. The model also accounts for potential "idiosyncratic risks" arising from severe ESG controversies. Through this comprehensive approach, Sustainalytics provides valuable insights into a company's ESG risk profile. (Sustainalytics, 2020)

4.4.1.2 Refinitiv ESG Score

Refinitiv has developed a comprehensive methodology for evaluating a company's environmental, social, and governance (ESG) performance in a transparent and objective manner. Refinitiv's ESG database stands as one of the industry's most extensive, encompassing more than 85% of the global market capitalization. Their assessment involves an extensive dataset comprising more than 630 company-level measures, with a focus on 186 measures that are deemed most comparable and material within each industry dating back to 2002. The selection of these measures considers various factors, including comparability, impact, data availability, and industry relevance, which may vary across different industry groups.

These measures are categorized into ten distinct categories, which contribute to the computation of three pillar scores and, ultimately, the overall ESG score. The final score is intended to reflect the company's ESG performance, commitment, and effectiveness based on publicly reported information. (Refinitiv, 2022)

4.4.2 Firm-level Sustainability Score

To construct fund sustainability measures, we begin by assembling the dataset on a stock level. As the data providers possess data and access limitations, we restricted the ESG score and fund-holding frequency to a yearly basis. After identifying the available stock-level ESG scores, we establish a linkage between the fund's holdings in our sample and the corresponding ESG scores. This is accomplished through the application of ISIN-code matching techniques, allowing us to obtain a distinct set of ESG scores specifically associated with the fund's holdings.

To ensure the comparability of ESG scores from different data providers, we employ yearly standardization by adjusting the scores to have a mean of zero and a standard deviation of one. It should be noted that higher values in the Refinitiv scores indicate better ESG performance, while in the case of Sustainalytics scores, higher values signify poorer ESG performance.

The matched sample used in our analysis is subject to missing data, as indicated in Appendix 1. Two noteworthy observations can be made from the table: a substantial increase in coverage in 2017 across both data agencies and a significant reduction in the mean of Sustainalytics' ESG risk scores in 2021. The limited availability of ESG ratings is a persistent challenge in ESG research, as highlighted in the existing literature. This trend of data unavailability is clearly visible in our data sample, where the number of ESG ratings per company more than triples over the ten-year period (see Appendix 2).

Similar to the observations made by Gibson et al. (2021), the restricted availability of ESG data poses a challenge in both the cross-section and time-series dimensions. To address this challenge, we adopt an approach of imputing the yearly mean for the missing weighted ESG scores under the assumption that unrated companies have ESG scores similar to the yearly mean of the sample.[6]

4.4.3 Fund-level Sustainability Score

The calculation of the fund ESG score is based on the normalized ESG score of the underlying stocks. However, one of the main challenges in this thesis pertains to the availability of sustainability ratings, particularly at the fund level, in the earlier part of the sample period. While Morningstar introduced its Sustainability Rating, encompassing fund and portfolio ratings, in 2016, fund-level sustainability ratings were not accessible prior to that date (Morningstar, 2021). Similarly, other popular data providers such as MSCI, Bloomberg, Refinitiv, and FTSE face similar limitations or provide sustainability ratings at the company level.

To address the issue of limited ESG coverage at the fund level, we follow the approach of Gibson and Krueger (2018) by computing the portfolio-level

[6] The mean is the average of all US-listed company ESG scores, calculated on an annual basis.

sustainability score. This involves utilizing the stock-level ESG scores from Refinitiv and Sustainalytics, along with the individual stock holdings per fund from Morningstar, to construct a replica of the mutual fund's portfolio and calculate the fund-level ESG score. This methodology allows for broader coverage and extends the sample period beyond the time-constrained globe ratings provided by Morningstar.

There is a concern regarding rating agencies making changes in their methodologies and criteria in the examination of sustainability rating that could have changed over time across providers. We deal with this issue by focusing on a relative measure, where the standardized ESG scores are normalized in ranks between 0 and 1. We estimate the Fund sustainability score by applying the following formula:

$$Fund\ Score_{j,t} = \sum_{i=1}^{Nj,t} w_{i,j,t} \times rk_t\left(Z - Score_{i,t}\right)$$

Where $Fund\ Score_{j,t}$ denotes the sustainability score of fund j in year *t*. $w_{i,j,t}$ is the equity value-weight of stock *i* in fund j at time *t*. $rk_t\left(Z - Score_{i,t}\right)$ represents the normalized rank of the standardized ESG score of stock *i* at time *t*. The standardized ESG scores are normalized in ranks between 0 and 1 and provide a relative position of the sustainability of stock *i* at time *t*. $N_{j,t}$ is the number of stocks in fund j at time *t*.

In the concluding stage of constructing the sustainability score, a rank-based classification is applied, partitioning the sustainability scores into distinct categories for each year. Specifically, funds attaining the highest scores are designated as top 10%, those within the intermediate range are assigned to the average group (80%) and funds with the lowest scores are categorized as low 10%. We carry out a robustness test considering the top 20%, average 60%, and bottom 20% levels to investigate whether the reported results are driven by a distinct cut-off level.

5.5 Potential Weaknesses in the Calculation of ESG Score

As discussed in the literature review, there exists considerable divergence among rating agencies regarding ESG assessments. In our analysis, we conduct separate analyses using data from each of the individual providers, which inherently relies on a single provider for the analysis. Consequently, the results may exhibit significant disparities. Despite our intention to adopt the approach proposed by Gibson and Krueger (2018) of utilizing a combined ESG score derived from averaging the scores of both providers, this methodology could not be implemented due to data access limitations. Such an approach is presumed to provide a more comprehensive representation of a firm's true sustainability and offer the opportunity to obtain the largest possible sample of company-level sustainability scores (Gibson, Krueger, Riand, & Schmidt, 2018).

The available data limitations posed challenges in obtaining ESG data at a monthly frequency, thereby potentially impacting the precision of our estimates, as our other variables are of monthly frequency. These limitations can be attributed to two main factors. Firstly, the maximum download capacity for fund holdings, considering the extensive number of holdings within funds that vary each quarter, limited our access to more data points.[7] Secondly, the absence of monthly ESG scores from both data providers further restricted our ability to analyze the finer dynamics of ESG and fund flows. Furthermore, the frequency of the data itself presents an additional issue. Given that our analysis covers a three-month period for both the COVID and conflict periods, relying solely on monthly data may result in an insufficient variation to effectively capture the nuances of these crises. To obtain a more precise understanding of the relationship between fund flows and ESG during crisis periods, it would be advantageous to obtain weekly data, enabling us to capture more granular fluctuations and better examine the dynamics at play.

When imputing the mean, a point of contention arises regarding whether to utilize an industry mean or the mean of the entire US market. It is argued that employing an industry mean is more appropriate because the ESG scores exhibit greater similarities across industries than within countries. The process of associating

[7] «More datapoints» refers to a longer sample period, and access to quarterly holdings data.

fund holdings with their corresponding ESG scores presents additional challenges in identifying industry per holding, further complicating data collection efforts.

Lastly, it is important to note that there is relatively low fund coverage in the first periods of our sample (see Appendix 1). For example, consider a scenario where one fund comprises a portfolio containing numerous companies with average ESG scores while another fund has a significant number of missing ESG scores. However, the true ESG score of the latter fund's portfolio is actually high. Due to the absence of ESG data for these companies, the second fund may receive an average ranking similar to that of the first fund. Consequently, the assumption of imputing the mean into missing ESG scores could potentially lead to erroneous interpretations of the actual ESG scores of the funds.

5.6 Net flow

Net flow is a measure of the movement of cash in and out of financial assets and can serve as an indicator of investor sentiment and behavior. The standard approach to calculating net flow in the literature is as follows:

$$\mathrm{FLOW}_{i,t} = \frac{TNA_t - TNA_{t-1}(1 + r_t)}{\mathrm{TNA}_{i,t-1}}$$

Where TNA_t is the end-of-month total net asset, TNA_{t-1} is the previous period total net asset, and r_t is the return in period t. This formula assumes that 100% of investors reinvest their distributions, which may not accurately reflect reality in most cases. This can lead to an underestimation of inflows, as asset growth is attributed to dividend reinvestments rather than to net new inflows (Morningstar, 2018). As a result, we use cash flows, and TNA already computed by Morningstar and aggregated at the fund level. Morningstar includes an additional step of adding back estimated distributions that have been cashed out.

$$Cash\ flow_t = TNA_t - TNA_{t-1}(1 + r_t) + \left(\left(\frac{TNA_{t-1}}{p_{t-1}}\right) * \sum_{i=1}^{t} d_i\right) * (1 - b)$$

Where d_i is the distribution (capital gain or dividend) during month t, p_{t-1} is the ending NAV in the previous period, and b is the reinvestment rate.

Using these components, we can calculate the net flow used throughout our analysis.

$$Net\ FLOW_{i,t} = \frac{Cash\ flow_t}{TNA_{i,t-1}}$$

5.7 Summary Statistics

Table 1: Summary Statistics full sample

This table display summary statistics for the entire period from January 2012 to December 2022. The variables Net flow, Std Dev, Monthly Return, and Expense are displayed as percentages. Net flow (Dollar), Size, and Age represent whole numbers, while Morningstar Rating takes on a value between 1 and 5. Panel A describes the variables in our sample. Panel B *presents the AUM-Weighted average over the full sample.*

Panel A: Variable Summary Statistics

Variable	Count	Mean	Std	Min	p25	p50	p75	Max
Net flow	137568	-0.441	2.999	-13.648	-1.248	-0.525	0.244	13.617
Net flow (Dollar)	137568	-9.318	77.620	-392.624	-15.006	-2.143	0.963	333.184
Size (Millions)	137568	3887.025	10251.324	12.470	216.810	844.995	2753.348	74759.646
Age (Months)	137568	288.741	163.629	36.000	192.000	252.000	329.000	1197.000
Std Dev	137568	14.553	7.711	2.910	9.278	12.731	18.072	42.032
Monthly Return	137568	1.006	4.516	-12.130	-1.314	1.352	3.621	13.009
Expense	137568	0.079	0.031	0.002	0.063	0.081	0.096	0.164
Morningstar Rating	137568	3.399	0.989	1.000	3.000	3.000	4.000	5.000

Panel B: AUM-Weighted Net flow

Variable	Count	Mean	Std	Min	p25	p50	p75	Max
Weigh. Net flow	130	-0.090	0.193	-0.688	-0.194	-0.088	0.021	0.660

Table 1 provides an overview of the data for the entire sample period from 2012 to 2022, with a focus on various variables, particularly the dependent variable "Net Flow." As shown in Panel A, the average "Net Flow" is slightly negative at -0.441, with a range from -13.648 to 13.617 and a median value of -0.525. It is interesting to note that the 25th percentile (-1.248) is significantly higher in absolute terms than the 75th percentile (0.244), indicating a predominance of negative net flow observations throughout the sample. This trend is consistent with the behavior of "Net flow (Dollar)", which represents the raw money inflows and outflows of the funds, with an average value of -9.318 million dollars over the entire sample, suggesting that investors, on average, withdraw more money than they invest.

The "Size" and "Age" variables stand out due to their higher values as whole numbers. To address these differences in metrics, we apply a logarithmic transformation in our regression analyses.

Furthermore, the equally weighted mean used for "Net Flow" may provide misleading information, as it treats small, non-growing funds the same as growing funds. To gain additional insights, we calculate the AUM-weighted net flow in Panel B. The "Weigh. Net flow" variable represents value-weighted net flows based on each fund's assets under management (AUM), giving smaller, non-growing funds a lower weight. The average of "Weigh. Net flow" remains negative at -0.09, confirming higher outflows than inflows for funds throughout the sample period.

Table 2: Average Variables for different ranks.

ESG Rank	Net flow	Size (Millions)	Age (Months)	Std Dev	Monthly Return	Expense	Morningstar Rating
Average ESG	-0.449	4107.543	289.661	14.354	1.020	0.077	3.383
Low ESG - Ref	-0.445	1573.946	258.016	16.656	0.992	0.093	3.449
High ESG - Ref	-0.418	5762.516	312.552	12.855	0.934	0.073	3.482
Low ESG - Sus	-0.334	1758.294	273.661	16.411	1.056	0.090	3.466
High ESG - Sus	-0.484	4719.485	307.012	13.397	0.932	0.077	3.457

This table presents the average Net flow, Size (Millions), Age (Months), Std Dev, Monthly Return, Expense, and Morningstar Rating for the different rankings over the entire period from January 2012 to December 2022. The Average ESG is funds that do not appear in any of the other ranks.

Table 2 presents the characteristics of the funds when divided into ESG ranks. We observe that both rankings indicating strong, sustainable performance, High ESG Refinitiv and Low ESG risk Sustainalytics, receive lower outflows compared to the other groups, with values of -0.418 and -0.334, respectively. Interestingly, the High ESG rankings from both providers exhibit significantly larger size and notably higher age compared to the Low rankings, which may be due to a positive correlation between size and age.

Table 3: Average Net flow of the ranks in Sub-Samples

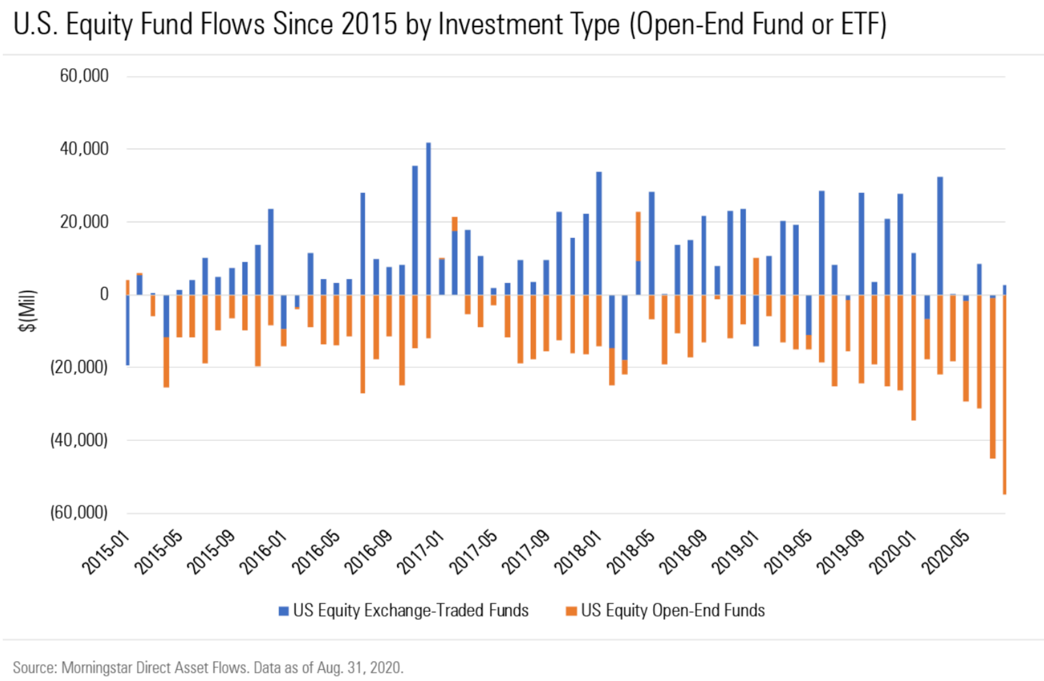
ESG Rank	Normal	Crisis
Average ESG	-0.437	-0.745
Low ESG - Ref	-0.428	-0.903
High ESG - Ref	-0.412	-0.605
Low ESG - Sus	-0.340	-0.298
High ESG - Sus	-0.466	-0.950

This table shows the average Net flow for the different ranks in both Normal and Crisis period

Based on the data presented in Table 3, it is evident that all ranks exhibit a negative average net flow during both the Normal and Crisis periods. Notably, the Low ESG risk category identified by Sustainalytics demonstrates the lowest outflow compared to all other rankings and the average ranked funds during Normal times, although the difference is relatively small. However, during the Crisis period, the Low ESG risk category stands out significantly with an average outflow of -0.298, which is considerably lower than all other rankings. Similarly, the High ESG category from Refinitiv also displays a lower outflow compared to the average ESG fund in crisis, with a value of -0.605. Both the ESG rankings, which indicate relatively poor performance in terms of ESG, the Low ESG score from Refinitiv, and the High ESG risk from Sustainalytics exhibit average outflows that are higher than the average fund, with values of -0.903 and -0.950, respectively.

Our analysis of the summary statistics reveals a negative average net flow, indicating that, on average, all funds in aggregate have experienced a higher outflow compared to inflow. We estimate that only approximately 60% of the funds in our sample have a higher total net asset (TNA) value at their last observation compared to their first observation. We hypothesize that this increase in TNA for 60% of the sample is attributable to returns, as the “Monthly Return” variable exhibits a higher positive value compared to the negative value of the “Net Flow” variable. Figure 1 illustrates the net flows of US Equity ETFs and US Equity Open-Ended funds over the period from January 1, 2015, to September 1, 2020. The graph shows a similar trend, with US Equity Open-Ended funds experiencing prevailing negative net flows from 2015 onwards, which encompasses a significant portion of our sample period.

Figure 1: Fund flows Equity Open-End Fund and ETF



The graph shows the US Equity fund flows for US equity open-ended and ETF funds from 2015-01-01 to 2020-09-01. From Morningstar. “U.S. Fund Flows Batter Equity Funds in August” by Thomas, T. & Grewal, S., 2020, https://www.morningstar.com/funds/us-fund-flows-batter-equity-funds-august
Copyright 2023 by Morningstar, Inc.

5.8 Diagnostic tests

To address potential issues within our sample, we conduct the test mentioned in section 4, the result from the Breusch-Pagan test (see Appendix 3), indicates the presence of heteroscedasticity. Furthermore, a Breusch-Godfrey test (see Appendix 4) reveals the existence of autocorrelation. To mitigate these problems, we employed Driscoll-Kraay standard errors throughout our analysis.

Additionally, after performing a variance inflation factor (VIF) analysis, we decided to exclude an interaction term between ESG rankings and Expense from our model due to high multicollinearity. This variable showed statistical insignificance, and its inclusion resulted in elevated VIF values for our primary variables of interest (Appendix 7), the ESG rankings. We made this decision to ensure the integrity and interpretability of our results, as recommended by Engqvist (2005).

Moreover, to determine the most appropriate regression model for our data, we conducted a restricted F-test (see Appendix 5), which led to the rejection of the null hypothesis. This suggests that a fixed effects (FE) model is more suitable than a pooled regression. Additionally, the Hausman test (see Appendix 6) also rejected the null hypothesis, indicating that an FE model is better suited for our data compared to a random effects (RE) model.

6. Results and Analysis

In the upcoming section, we will present the empirical findings related to the fund flow performance of high-ranked ESG funds and low-ranked ESG funds. These results, which address our research questions, will be presented, and visualized in the subsequent chapter, building upon the methodology previously discussed. This study examines the impact of ESG factors on fund flows throughout the entire sample period, including normal and crisis periods, with a specific focus on the effects of the COVID-19 crisis and the Russia-Ukraine conflict.

6.1 Fixed Effects Model

6.1.1 Full sample Fixed Effects Model Results

Table 4 presents the results of a fixed effects model applied to the entire sample. In this model, both ESG scores and ESG rankings from Refinitiv and Sustainalytics are included as independent variables separately. We emphasize that strong performance in relation to ESG is observed in funds with a High ESG rank according to Refinitiv and funds with a Low ESG risk rank according to Sustainalytics.

Table 4: **Baseline Regression**

	(1)	(2)	(3)	(4)	(5)
Dep. Variable	Net flow	Net flow	Net flow	Net flow	Net flow
No. Observations	137568	137568	137568	137568	137568
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0596	0.0601	0.0597	0.0599	0.0599
F-statistic	1441.4	1090.7	1082.0	868.03	868.81
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000	0.0000
Monthly Return	0.0997***	0.1162***	0.1082***	0.0942***	0.0938***
	(11.027)	(9.9303)	(7.8515)	(10.013)	(9.8645)
Morningstar Rating	0.9266***	0.9236***	0.9261***	0.9269***	0.9261***
	(21.712)	(25.152)	(25.027)	(25.253)	(25.144)
Expense	-3.4707**	-3.3594**	-3.4444**	-3.4451**	-3.4229**
	(-2.0542)	(-2.1388)	(-2.1916)	(-2.1872)	(-2.1698)
Std Dev	-0.0179**	-0.0195***	-0.0185**	-0.0179**	-0.0184**
	(-2.3365)	(-2.6221)	(-2.5043)	(-2.4482)	(-2.4822)
Log Age	-1.1076***	-1.1192***	-1.0957***	-1.1034***	-1.0915***
	(-4.8502)	(-5.4853)	(-5.3478)	(-5.4063)	(-5.3446)
Log Size	-0.6908***	-0.6833***	-0.6914***	-0.6916***	-0.6921***
	(-17.146)	(-18.091)	(-18.485)	(-18.595)	(-18.684)
ESG Score		-1.2523***	-0.1416		
		(-4.1186)	(-0.8670)		
ESG * M. Return		-0.0375**	-0.0213		
		(-2.4185)	(-1.0059)		
Low ESG				-0.1019*	0.0298
				(-1.7547)	(0.6196)
High ESG				-0.0093	-0.0982**
				(-0.2650)	(-2.5657)
Low ESG x M. Return				0.0231***	0.0253***
				(3.5197)	(3.2227)
High ESG x M. Return				0.0164**	0.0120*
				(2.3229)	(1.7835)
FE	Yes	Yes	Yes	Yes	Yes

The table presents the results from regressions of Net Flow on different ESG scores and rankings. Columns (1) includes only control variables, column (2) is the ESG score from Refinitiv, and column (3) is the ESG risk score from Sustainalytics, while (4) and (5) rank the funds into top and bottom 10% for Refinitiv and Sustainalytics, respectively. Driscoll-Kraay standard errors are used across every regression to account for heteroscedasticity and autocorrelation; t-statistics are in parenthesis. *, **, *and* ****indicates significance at 10%, 5% and 1%, respectively.*

In Column (1), we include only the control variables without the variables of interest, as it allows for a focused examination of the individual effects of the control variables on the dependent variable. By isolating these effects, it helps establish a baseline understanding of their impact before introducing the variables of interest, providing a clearer picture of the influence of the control variables on the outcome. The results indicate that a one-unit increase in Monthly Return corresponds to a 0.0997% higher inflow. Similarly, Morningstar star ratings have a positive effect on net flows, with a one-unit increase in the Morningstar rating leading to a 0.9266% inflow. Fund expense ratio and fund volatility demonstrate a negative impact on fund flows, with a one-unit increase resulting in -3.4707% and -0.0179% outflows, respectively. These regressors are significant at the 0.05 level, while the remaining control variables show significance at the 0.01 level. Finally, the results indicate that a one-unit increase in both the logarithm of fund age and the logarithm of size leads to outflows of -1.1076% and -0.6908%, respectively.[8]

[8] The application of logarithmic transformation to age and size variables results in coefficients that appear relatively large in magnitude.

Columns (2) and (3) introduce fund ESG scores based on Refinitiv ESG scores and Sustainalytics ESG risk scores. Additionally, the interaction term between fund ESG score and return is included. The results in Column (2) show negative and statistically significant coefficients for the ESG score and the interaction term. A one-unit increase in the Refinitiv fund ESG score implies an outflow of -1.2523, with a t-statistic of -4.1186.[9]

Columns (4) and (5) of the table introduce ESG rankings, dividing them into top 10% scores (High ESG) and bottom 10% scores (Low ESG) based on ESG scores from Refinitiv and Sustainalytics. The Low ESG rank, according to Refinitiv, shows an outflow of -0.1019%, while funds with High ESG risk, according to Sustainalytics, exhibit an outflow of -0.0982%. Remarkably, the only rank that demonstrates statistical significance at the 5% level throughout the observation period is the high ESG risk category, suggesting its impact on fund outflows. Furthermore, the table shows a positive interaction effect between fund ESG rank and monthly return for both high and low ranks in Columns (4) and (5).

6.1.2 Sub-sample Fixed Effects Model Results

Table 5 displays the results of the fixed effects model conducted on two distinct sub-samples, namely the normal and crisis periods. The model employed in this analysis is identical to the one conducted in section 6.1.1, with the exception that the ESG score variable has been excluded from the model.

[9] Note that the fund ESG scores are normalized, meaning that the estimated coefficient represents a -1.2523% decrease in net flow for a standard-deviation increase in the normalized fund score. Hence a one-unit increase equals a one-standard-deviation increase.

Table 5: **Rankings in Normal and Crisis periods.**

	Normal		Crisis	
	(1)	**(2)**	**(3)**	**(4)**
Dep. Variable	Net flow	Net flow	Net flow	Net flow
No. Observations	132221	132221	6422	6422
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0601	0.0601	0.0443	0.0482
F-statistic	837.94	838.34	24.690	26.964
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000
Low ESG	-0.0932	0.0088	-0.3348**	0.3476***
	(-1.6381)	(0.1863)	(-2.3211)	(2.7496)
High ESG	-0.0127	-0.0848**	-0.1071	-0.6105***
	(-0.3343)	(-2.1469)	(-0.5187)	(-4.7198)
Monthly Return	0.0961***	0.0956***	0.0336	0.0213
	(9.8906)	(9.6591)	(0.6543)	(0.4797)
Morningstar Rating	0.9300***	0.9295***	0.5652***	0.5597***
	(24.721)	(24.621)	(7.0590)	(8.5137)
Log Size	-0.6946***	-0.6949***	-2.3325***	-2.3032***
	(-18.079)	(-18.136)	(-6.1101)	(-6.1594)
Expense	-3.7921**	-3.7834**	3.2501	3.3209
	(-2.3708)	(-2.3619)	(0.5143)	(0.5843)
Std Dev	-0.0162**	-0.0165**	-0.0527***	-0.0615***
	(-2.2654)	(-2.2847)	(-2.7404)	(-2.9735)
Log Age	-1.0536***	-1.0439***	-2.7025*	-2.0099
	(-5.0998)	(-5.0469)	(-1.7070)	(-1.3401)
Low ESG x M. Return	0.0219***	0.0249***	0.0242***	0.0447***
	(3.0408)	(3.0383)	(3.5554)	(7.3210)
High ESG x M. Return	0.0178**	0.0118	0.0110	-0.0003
	(2.0493)	(1.4621)	(0.7093)	(-0.0315)
FE	Yes	Yes	Yes	Yes

*The table presents the results from regressions of Net Flow on Low, and High ESG ranks in normal times. Columns (1) and (2) is ranks created using ESG scores from Refinitiv, while columns (3) and (4) uses ESG risk score from Sustainalytics. Driscoll-Kraay standard errors are used across every regression to account for heteroscedasticity and autocorrelation, t-statistics are in parenthesis. *, **, and *** indicates significance at 10%, 5% and 1%, respectively.*

Columns (1) and (2) present the results for High and Low ESG ranks in the normal sub-sample using Refinitiv and Sustainalytics data, respectively. The analysis reveals that the only significant rank in the normal sub-sample is the High ESG rank, according to Sustainalytics. Funds with a High ESG rank experience an outflow of -0.0848% compared to funds with an average ESG rank, remaining statistically significant at a 5% level. Columns (3) and (4) present the results for the crisis sub-sample using Refinitiv data. Analyzing the rankings from Refinitiv, we find that funds with a Low ESG rank exhibit a significant outflow of -0.3348% compared to the average fund. Funds with a High ESG rank also show a negative coefficient, although it is not statistically significant. For the rankings provided by Sustainalytics ESG risk, we find statistical significance for both the High ESG risk rank and Low ESG risk rank during crisis periods. The t-stats for high rank and low rank are -4.7918 and -2.7496, respectively, with coefficients of -0.6015 and 0.3476.

6.1.3 Fixed Effects Model with Dummies

Table 6 shows the results of the fixed effects model conducted with dummy variables for the Covid-19 and Russia-Ukraine conflict.

***Table 6*: ESG Ranks in Different Crisis**

	(1)	(2)
Dep. Variable	Net flow	Net flow
No. Observations	137568	137568
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0599	0.0600
F-statistic	620.36	621.72
P-value (F-stat)	0.0000	0.0000
Low ESG	-0.0978*	0.0079
	(-1.6505)	(0.1650)
High ESG	-0.0105	-0.0812**
	(-0.2757)	(-2.1487)
Monthly Return	0.0943***	0.0935***
	(10.040)	(9.8202)
Morningstar Rating	0.9271***	0.9259***
	(25.209)	(25.150)
Log Size	-0.6917***	-0.6925***
	(-18.596)	(-18.648)
Expense Ratio	-3.4482**	-3.4189**
	(-2.1871)	(-2.1696)
Std Dev	-0.0179**	-0.0185**
	(-2.4458)	(-2.5079)
Log Age	-1.1031***	-1.0930***
	(-5.4010)	(-5.3462)
Low ESG x M. Return	0.0225***	0.0278***
	(3.5322)	(3.4206)
High ESG x M. Return	0.0162**	0.0093
	(2.1406)	(1.3297)
Low ESG x Covid	-0.1748*	0.4255**
	(-1.7059)	(2.2046)
High ESG x Covid	-0.2413***	-0.2874**
	(-3.1030)	(-2.4460)
Low ESG x Conflict	0.0003	0.2596*
	(0.0035)	(1.9362)
High ESG x Conflict	0.2133	-0.2712***
	(1.3325)	(-4.0401)
FE	Yes	Yes

The table presents the results from regressions of Net Flow on the different rankings with dummies for both Covid and Conflict. Columns (1) is based on the ESG score from Refinitiv, while (2) is the ESG risk score from Sustainalytics. Driscoll-Kraay standard errors are used across every regression to account for heteroscedasticity and autocorrelation, t-statistics are in parenthesis. *, **, and ****indicates significance at 10%, 5% and 1%, respectively.*

In column (1), the results based on ESG rankings from Refinitiv reveal a weak level of significance at the 10% level for funds with a Low ESG rank, indicated by a coefficient of -0.0978. Furthermore, the significance of these rankings is only observed during the Covid period, not the Conflict period. In the Covid period, funds with a Low ESG rank experienced an outflow of -0.1748%, while funds with a High ESG rank faced an even higher outflow of -0.2413% compared to funds with an average ESG rank. Moving on to column (2) and considering the Sustainalytics rankings, a similar trend is observed for funds with poor sustainability performance. Funds with a High ESG rank experience an outflow of -0.0812%. During the Covid-19 crisis, both high ESG and low ESG ranks demonstrate statistical significance at the 5% level, with coefficients of -0.2874 and 0.4225, respectively. In the Conflict period, the coefficients are -0.2712 and 0.2596, with statistical significance at the

1% and 10% levels. We observe a reduction of 0.1659 between the coefficients of Low ESG risk in Covid compared to Conflict.

6.2. Robustness analysis

To test the robustness of our result, we test our models according to the modification mentioned in section 5.3.

6.2.1 Fund family fixed effects

When analyzing the results of fund family fixed effects (Appendix 9), the results are consistent with the rankings provided by Sustainalytics ESG risk. However, the same cannot be said about the rankings provided by Refinitiv ESG, which yield some inconsistent results.

The previously observed negative significance of the Refinitiv ESG score in our findings becomes statistically insignificant, while Sustainalytics retains its weak statistical significance. Regarding the rankings, the High ESG risk from Sustainalytics continues to exhibit signs as the sole rank within the entire sample, although it is noteworthy that the impact has diminished to -0.0532 with a lower absolute t-stat of 1.6561. Upon analyzing the sub-samples, it is evident that the rankings from Sustainalytics are the only ones indicating an effect on the net flow in our crisis sub-sample. In this regard, the Low ESG Sustainalytics ranking continues to demonstrate a positive net flow in comparison to the average fund, with a value of 0.3818. The High ESG risk ranking experiences a negative net flow of -0.3365, which, despite being a lesser outflow than initially observed, remains negative.

In the final model, it is evident that the initial effects attributed to the rankings outside of the Covid and Conflict periods become statistically insignificant. For ESG rankings from Refinitiv in Covid, High ESG remains to contribute negatively to net flow, while the initial negative effect from Low ESG becomes statistically insignificant. These rankings remain insignificant in Conflict as well. The rankings provided by Sustainalytics remain consistent with the previous results, with minor differences.

6.2.2 Wider ESG rankings

The impact of rankings over the entire period is inconsistent, with Sustainalytics' ESG ranks showing some consistent tendencies during market downturns, while Refinitiv's rankings produce inconclusive results. Overall, neither Refinitiv nor Sustainalytics rankings consistently affect net flows throughout the entire period.

In the sub-sample analyses, Refinitiv High ESG ranks become significant in both the Normal period and Crisis period, with higher outflows. Previously significant rankings for Low ESG Refinitiv in Crisis and High ESG risk Sustainalytics in the normal period become insignificant. Sustainalytics' ESG risk rankings continue to exhibit a consistent effect during the Crisis period, although High ESG risk funds experience lower outflows compared to our main findings. Nonetheless, investors tend to invest more in Low ESG-risk funds and withdraw more from High ESG-risk funds during market downturns.

In the separate analyses of Covid and Conflict, Refinitiv's ESG rankings show a shift in coefficients for both Low ESG and High ESG funds, where only Low ESG is statistically significant. Sustainalytics' ESG risk rankings remain relatively consistent, with Low ESG risk funds experiencing positive inflows during Covid and High ESG risk funds showing higher outflows during Conflict. However, the High ESG risk in Covid and Low ESG risk in Conflict becomes statistically insignificant.

6.3 Discussion of empirical results

In this section, we aim to provide a discussion of our results within the context of the research questions posed in this study. Additionally, we explore potential reasons that may explain why the observed outcomes deviate from the expectations set forth by previous literature.

Hypothesis 1: Does fund ESG rank affect the net flows of US mutual equity fund?

Across the conducted regressions in Tables 4, 5, and 6, a consistent pattern emerges, indicating that open-end equity mutual funds exhibiting poor sustainability performance experience higher outflows compared to the average fund. The High ESG risk rankings provided by Sustainalytics exhibit statistical significance during the entire sample period and normal period but do not exhibit consistency in the robustness test.[10] Conversely, the Low ESG rankings from Refinitiv demonstrate weaker significance and weak robustness. Regarding funds with strong sustainability performance, the result does not reveal a significant effect on net flow over the entire period. The fund-level ESG scores exhibit an unexpected effect, showing a negative relationship with net flows based on data from Refinitiv, although it is not robust. However, the Sustainalytics fund ESG score remains statistically insignificant in this regard.

Our findings indicate that mutual fund investors collectively do not prefer funds with poor sustainability performance, as they tend to withdraw a larger portion of their investments from such funds. This is shown by the High ESG risk from Sustainalytics having statistically significant outflows higher than the average ESG fund. These findings display a certain level of correspondence with the research of Hartzmark and Sussman (2019) and Ammann et al. (2018), where investors allocate less capital to funds with unfavorable sustainability ratings. Despite differences in the sustainability measures and methods used in these studies compared to our methodology, their results provide a satisfying view of sustainable mutual fund flow applicable to our analysis. However, it is important to acknowledge that the robustness tests conducted on the fund ESG rankings from both providers yielded less conclusive results. This poses a challenge in validating our empirical findings.

Given that funds with strong sustainability performance in our analysis remain statistically insignificant across both Refinitiv and Sustainalytics rankings, it is not possible to draw conclusive implications regarding the funds exhibiting strong sustainability performance on fund flows. Consequently, these findings do not provide evidence to support findings that investors allocate more investments in highly sustainable funds or that investors treat sustainability as a positive fund attribute (Hartzmark and Sussman, 2019; Ammann et al., 2018). The same applies to the arguments made by Bollen (2007) and Renneboog et al. (2008) regarding investors' consideration of non-financial attributes, including sustainability, in their investment decisions, which are not corroborated by our findings. Thus, our

[10] High ESG risk rank show significance in the robustness test changing to fund family fixed effects. In the other robustness test, changing the cut-off to 20/60/20 reverse the significance of the rankings, making high ESG risk rank and low ESG rank insignificant, while low ESG risk rank and high ESG rank significant.

analysis does not provide statistical evidence for the claim that mutual fund investors generally exhibit a preference for funds with high sustainability rankings.

Analyzing the fund ESG scores, we find that the Refinitiv ESG score reveals a negative relationship with the net flow, although not robust. This finding contradicts our initial beliefs, although it is reasonable to attribute it to the potential weaknesses in the dataset. There has been a significant increase in ESG coverage after 2017, and the limited coverage prior to this period may distort the impact of the fund's ESG score. Additionally, the impact of the fund's ESG score on net flow is relatively small, as a one standard deviation increase is uncommon due to the narrow range of the fund ESG scores. The fund ESG score based on Sustainalytics is statistically insignificant, which might indicate that the ESG score itself may not be the sole determinant of fund flows, as investors may consider other factors or have differing interpretations of ESG ratings.

Considering the null hypothesis, our results initially indicate that funds with poor sustainability performance experience higher outflows compared to the average ESG fund. However, due to the low significance and small magnitude of the coefficient, as well as the changes observed after conducting robustness tests, we lack sufficient evidence to reject the null hypothesis. This implies that funds with both high and low sustainability performance do not exhibit significantly different net flows compared to the average funds in our sample.

Hypothesis 2: Does high-ranking ESG funds and low-ranking ESG funds experience differences in net flows compared to average ESG funds in market downturns?

The findings presented in Table 5 highlight the statistical significance of the High and Low ESG rankings reported by Sustainalytics within the Crisis sub-sample. Importantly, these results maintain their robustness even after conducting additional tests. In contrast, our analysis did not reveal the same level of significance for the ESG rankings generated by Refinitiv ESG. These rankings lacked both consistency and statistical significance in our study. This suggests that investors do not exhibit a clear preference for funds based on Refinitiv ESG rankings during crisis periods, in contrast to the strong preference observed for funds with strong sustainability performance based on Sustainalytics rankings.

Our analysis of the ESG risk rankings provided by Sustainalytics indicates that investors have a clear preference for funds with strong sustainability performance and show less interest in funds with weak sustainability performance during crisis periods. This is consistent with our hypothesis, based on the studies of Pastor and Vorsatz (2020) and Ferriani and Natoli (2020), that funds with strong sustainability performance would receive higher net flows during market downturns. Our results confirm that sustainable funds do indeed receive higher net flows compared to the average fund, indicating a greater positive impact on fund flows compared to non-sustainable funds.

Furthermore, our findings align with additional findings of Pastor and Vorsatz (2020), which suggests that sustainability is perceived as a fundamental requirement rather than a discretionary luxury. We observe that funds with Low ESG risk experience lower outflows compared to funds with High ESG risk during crises, as evidenced in Table 3. Additionally, our regression analysis in Table 6.2 demonstrates that funds with Low ESG risk attract positive net flows, while those with high ESG risk experience negative net flows. These results are consistent with the "flight to safety" effect identified by Ferriani and Natoli (2020), wherein funds with high ESG risk rankings experience significant outflows. It is important to acknowledge that there are differences in methodology, sustainability measures, and the specific crisis periods defined in these studies. However, their findings provide a compelling perspective on the flow of sustainable mutual funds, which is applicable and supportive of our own analysis.

Contrary to our previous discussion under hypothesis 1, our findings align with and further extend the conclusions of Hartzmark and Sussman (2019). Our results indicate that investors indeed exhibit a tendency to allocate their funds towards highly sustainable funds while withdrawing from low sustainable funds during market downturns. This suggests that sustainability is perceived as a positive attribute by investors, particularly in challenging market conditions. By corroborating the findings of Hartzmark and Sussman, our study provides additional evidence supporting the notion that sustainability plays a significant role in investment decision-making during periods of market downturns.

Our dataset reveals a significant disparity among ESG providers, in line with previous research by Berg et al. (2019) and Gibson et al. (2021), which highlights the considerable variation in ESG ratings across different providers. Our results show a recurring pattern of insignificant ESG fund rankings and scores, particularly during normal periods and throughout the entire sample period. Additionally, there is a high degree of difference between the ESG rankings. As discussed in section 5.5, factors such as our fund-level ESG calculation could affect the significance of the variables. Our analysis shows that the ESG rankings based on Sustainalytics are more significant, which may be due to the fact that this metric is used to calculate the Morningstar Sustainability Globes, a measure that has been found to be a driver for fund flows in previous literature (Hartzmark and Sussman, 2019). This is likely due to the easy accessibility of this measure compared to creating your own ESG from the underlying portfolios of funds, as is done with the Refinitiv ESG score.

Considering the null hypothesis, our results suggest that funds with strong sustainability performance experience higher inflows, and funds with poor sustainability performance experience higher outflows compared to funds with average ESG rankings during crisis periods. As high ESG risk rank and low ESG risk rank exhibit statistical significance and robustness, it allows us to reject the null hypothesis. Hence, funds with high or low ESG rankings have a significantly different fund flow compared to funds with average ESG rankings.

Hypothesis 3: Does high-ranking ESG funds and low-ranking ESG funds experience differences in net flows compared to average ESG funds in Covid-19 and Russia-Ukraine conflict?

Upon examining the findings presented in Table 6, a consistent trend emerges in accordance with our discussion in Hypothesis 2. Specifically, the High and Low ESG rankings provided by Sustainalytics during the COVID-19 crisis and Russia-Ukraine Conflict exhibit statistical significance but with limited robustness. In contrast, the rankings derived from Refinitiv demonstrate significance solely during the COVID-19 crisis, lacking robustness. Moreover, noteworthy disparities in the coefficients between the two crises are observed.

Our analysis, using Sustainalytics data, indicates a preference among fund investors for funds demonstrating strong sustainability performance during crises, similar to our previous findings. This aligns with the conclusions drawn by Pastor and Vorsatz (2020), Ferriani and Natoli (2020), and Hartzmark and Sussman (2019), providing evidence that sustainable funds attract more flows than the average fund, while non-sustainable funds experience lower flows during the Covid and Conflict periods.

Furthermore, the results provide some indication that inflows into low ESG risk funds are lower during the Russia-Ukraine Conflict compared to the COVID-19 period. However, the outflows from high ESG risk funds are similar in both crises. The estimates lack robustness for both crisis periods, and we cannot confidently conclude differences in mutual fund flows between high- and low-ranking funds during the crises, as no statistical test was performed on the significant difference.

Considering the null hypothesis, our results suggest similar findings to that of hypothesis 2 based on ESG data from Sustainalytics. Mutual funds with strong sustainability performance attract higher inflows, while funds with poor sustainability performance experience higher outflows, in comparison to funds with average ESG rankings during the periods of COVID-19 and the Russia-Ukraine Conflict. Even if both statistical significance and robustness vary in the results, Low ESG rank during COVID-19 and High ESG rank during Conflict are robust and statistically significant, from Sustainalytics. This implies that funds with high or low ESG rankings have a significantly different fund flow compared to funds with average rankings during the crises. Hence, we reject the null hypothesis.

7 Conclusion

This paper investigates the relationship between fund ESG and net flows of US mutual equity funds. Using data from Morningstar Direct and Eikon Refinitiv, we present a compelling argument that mutual funds demonstrating strong sustainability performance tend to attract larger inflows, whereas funds displaying weak sustainability performance experience greater outflows during market downturns. However, our findings do not indicate any significant impact of sustainability performance on fund flows in general.

Our findings support the findings of Pastor and Vorsatz (2020) and suggest an extension to the findings of Hartzmark and Sussman (2019) and Ammann et al. (2018). Furthermore, our findings are consistent with the concept of a "flight to safety," where investors allocate their investments towards perceived secure options during times of stress, suggesting a preference for highly sustainable funds during crises and a preference against non-sustainable funds (See, e.g., Ferriani and Natoli (2020), and Parida and Wang (2008)).

We acknowledge potential limitations in our methodology for computing fund-level ESG scores, particularly in relation to the approach proposed by Gibson and Krueger (2018). Due to data access limitations, we were unable to combine scores from different providers as intended, which hinders the generation of a comprehensive ESG estimate reflecting a "true sustainability score." Moreover, factors such as a significant number of missing ESG scores prior to 2017, limitations in data frequency, and the uniqueness of our method in utilizing ESG scores dating back to a time before ESG became a commonly used term in finance, raise concerns about the consistency and explanatory power of our ESG estimates. While various measures of sustainability are found in the literature, our specific method deviates from commonly employed approaches.[11]

Further research is needed to deepen our understanding of these dynamics and their implications for sustainable investing in different contexts. Future research in sustainable investing can focus on examining the impact of ESG factors on non-equity funds, exploring different markets outside the United States, and

[11] See, e.g., Ghoul and Karoui (2017), utilizing a similar method calculating sustainability score. However, they use CSR as a basis.

investigating investor preferences for sustainable funds in the aftermath of environmental disasters. These areas of study would provide a more comprehensive understanding of the relationship between ESG and investor behavior, expand the analysis to different asset classes and global contexts, and explore the role of sustainability in addressing environmental challenges. Such research can contribute valuable insights for investors, policymakers, and industry practitioners aiming to integrate ESG considerations into their investment decisions.

Bibliography

Ammann, M., Bauer, C., Fischer, S., Muller, P., (2018). The Impact of the Morningstar Sustainability Rating on Mutual Fund Flows. *University of St.Gallen School of Finance Research Paper Series*. https://doi.org/10.1111/eufm.12181

Baker, S. R., N., Bloom, S. J., Davis, K. J. Kost, M. C. Sammon, and T. Viratyosin, (2020). The Unprecedented Stock Market Reaction to COVID-19. *The Review of Asset Pricing Studies 10*. Vol. (4), pp. 742–758. http://www.nber.org/papers/w26945

Baumol, William J. and Wallace E. Oates, 1979, Economics, Environmental Policy, and Quality of Life, Inglewood Cliffs, NJ: Prentice-Hall.

Barko, T., M. Cremers, and L. Renneboog (2021). Shareholder engagement on environmental, social, and governance performance. Journal of Business Ethics. https://link.springer.com/article/10.1007/s10551-021-04850-z

Benson, K., L., Humphrey, J., E., (2008). Socially Responsible Investment Funds: Investor Reaction to Current and Past Returns. *Journal of Banking and Finance, Forthcoming*.

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

Bloomberg, Kishan, S., (2022). ESG by the Numbers: Sustainable Investing Set Records in 2021. Green Finance. https://www.bloomberg.com/news/articles/2022-02-03/esg-by-the-numbers-sustainable-investing-set-records-in-2021#xj4y7vzkg

Bollen, P. B. (2007). Mutual Fund Attributes and Investor Behavior. Journal of Financial and Quantitative Analysis. 42, 683-708. https://www-cambridge-org.ezproxy.library.bi.no/core/services/aop-cambridge-core/content/view/7F2F572F662691436869D481C3DFA7FC/S0022109000004142a.pdf/mutual-fund-attributes-and-investor-behavior.pdf

Brooks, C. (2019). Introductory econometrics for finance (4th ed.). Cambridge University Press.

Dottling, R., & Kim, S., (2022). Sustainability Preferences Under Stress: Evidence from COVID-19. Journal of Financial and Quantitative Analysis, Forthcoming. pp. 1-39 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3656756

Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. Review of Economics and Statistics, 80(4), 549-560. https://doi.org/10.1162/003465398557825

Engqvist, L. (2005). *The mistreatment of covariate interaction terms in linear model analyses of behavioral and evolutionary ecology studies*. Animal Behaviour, 70(4), 967-971. https://doi.org/10.1016/j.anbehav.2005.01.016

Ferriani, F., and Natoli, F., (2020). ESG risks in times of Covid-19. *Applied Economics Letters*. Vol. 28, pp. 1537-1541.

Gibson, R. and Krueger, P., 2018, The sustainability footprint of institutional investors, Swiss Finance Institute working paper.

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

Greene, W. H. (2011). Econometric analysis (7th ed.). Prentice Hall.

Guercio, D., & Tkac, P. (2000). The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds Versus Pension Funds. FRB Atlanta Working Paper 2000-21. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=166431

Guercio, D., & Tkac, P. (2008). Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow. Journal of Financial and Quantitative Analysis, 43(4), 907-936. doi:10.1017/S0022109000014393

Hartzmark, S. M., & Sussman, A. B. (2019). Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. Journal of Finance. 74, 2789-2837. https://onlinelibrary-wiley-com.ezproxy.library.bi.no/doi/10.1111/jofi.12841

Humphrey et al., (2016). What is Different about Socially Responsible Funds? A Holdings-Based Analysis. *Journal of Business Ethics*. Vol. 138, pp. 263-277. https://link.springer.com/article/10.1007/s10551-015-2583-7

Izzeldin, M., Muradoglu, Y., G., Pappas, V., Petropoulou, A., Sivaprasad, S., (2023). The Impact of the Russia-Ukrainian War on Global Financial Markets. *International Review of Financial Analysis*. Vol. 87.

Jones, S., Laan, S., Frost, G., Loftus, J., (2008). The Investment Performance of Socially Responsible Investment Funds in Australia. *Journal of Business Ethics*. Vol. 80, pp. 181-203.

Kempf, A., & Osthoff, P. (2008). SRI Funds, Nomen est Omen. Journal of Business, Finance & Accounting, 35, 1276–1294. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1468-5957.2008.02107.x

Markowitz, H. (1952) Portfolio Selection. Journal of Finance. 7, 77–91.

McKinsey, (2022). Does ESG Really Matter – and Why. McKinsey Quarterly.

https://www.mckinsey.com/capabilities/sustainability/our-insights/does-esg-really-matter-and-why

Morgan Stanley, (2020). Morgan Stanley Sustainable Signals: Asset Owners See Sustainability as Core to Future of Investing. https://www.morganstanley.com/press-releases/morgan-stanley-sustainable-signals--asset-owners-see-sustainabil

Morgan Stanley, (2023). Despite market challenges, demand for sustainable funds remains strong. Institute for sustainable investing. https://www.morganstanley.com/ideas/sustainable-funds-performance-demand

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

Morningstar Research. (2018, August 31). Estimated Net Cash Flow Methodology. Morningstar. https://www.morningstar.com/content/dam/marketing/shared/research/methodology/765555_Estimated_Net_Cash_Flow_Methodology.pdf

MSCI, (n.d.). The Evolution of ESG Investing. https://www.msci.com/esg-101-what-is-esg/evolution-of-esg-investing

MSCI, (2023). Global Markets One Year After Russia's Invasion of Ukraine. https://www.msci.com/www/blog-posts/global-markets-one-year-after/03668219477

Nofsinger, J., Varma, A. (2014). Socially responsible funds and market crises. Journal of Banking and Finance. 48, 180-193. https://www.sciencedirect.com/science/article/pii/S0378426613004883?via%3Dihub

Oikonomou, I., Brooks, B., Pavelin, S. (2012). The impact of corporate social performance on financial risk and utility: a longitudinal analysis. Financial Management. 41, 483-515.

Parida, S., and Wang, Z., (2018). Financial Crisis and Corporate Social Responsible Mutual Fund Flows. *International Journal of Financial Studies*.

Pastor, L. and B. Vorsatz (2020). Mutual fund performance and flows during the COVID-19 crisis. The Review of Asset Pricing Studies 10 (4), 791–833. https://www.nber.org/system/files/working_papers/w27551/w27551.pdf

Renneboog, L., Horst, J. T., Zang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. Journal of Banking and Finance. 32, 1723-1742.

Renneboog, L., Ter Horst, J., Zhang, C., (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*. Vol 20, pp. 562-588.

Refinitiv. (2022). Refinitiv ESG scores methodology. https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf

Sirri, E., & Tufano, P., (1998). Costly search and mutual fund flows. Journal of Finance 53, 1589–1622. https://doi.org/10.1111/0022-1082.00066

Sustainalytics ESG score rating 2020: https://connect.sustainalytics.com/hubfs/SFS/Sustainalytics%20ESG%20Risk%20Ratings_Issuer%20Backgrounder.pdf

Thomas, T., & Grewal, S. (2020). U.S. Fund Flows Batter Equity Funds in August. Morningstar. https://www.morningstar.com/funds/us-fund-flows-batter-equity-funds-august

United Nations, (2004). Who Cares Wins: Connecting Financial Markets to a Changing World. https://www.unepfi.org/fileadmin/events/2004/stocks/who_cares_wins_global_compact_2004.pdf

Warther, VA. (1995). Mutual Fund Flows and Security Returns. Journal of Financial Economics. 39, Pages 209-235. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5453

World Health Organization. (n.d.). COVID-19 situation in the WHO European Region. https://www.who.int/europe/emergencies/situations/covid-19

Wooldridge, J. M. (2019). Introductory Econometrics: A Modern Approach (7th ed.). Cengage Learning.

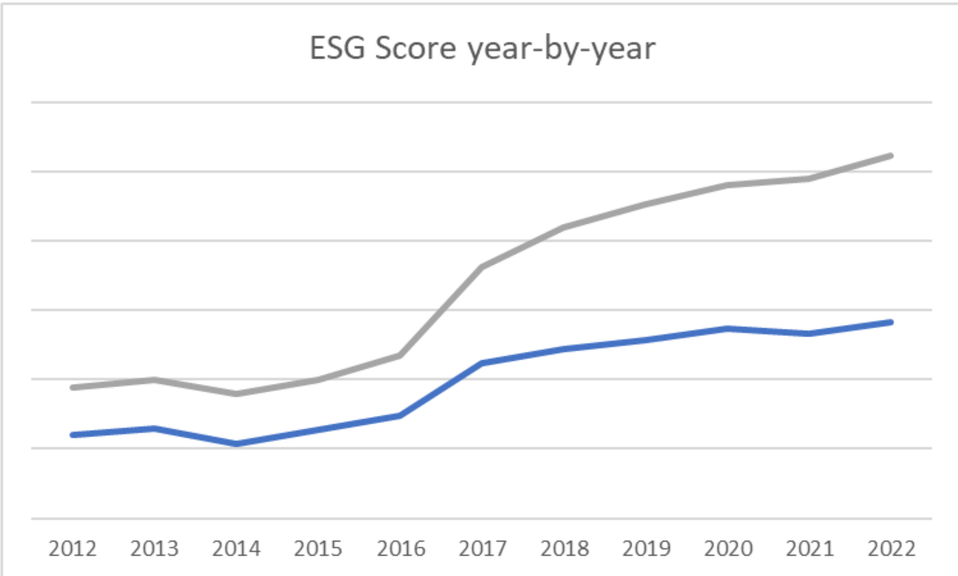
Appendices

Appendix 1: **Missing ESG per Year**

	REFINITIV		**SUSTAINALYTICS**	
DATE	Missing ESG score – Fund Level	Mean	Missing ESG score – Fund Level	Mean
2012	34.93%	46.60	35.23%	56.60
2013	34.16%	46.92	34.37%	56.47
2014	34.20%	47.35	41.40%	58.05
2015	33.61%	47.52	38.30%	58.99
2016	28.65%	45.66	34.36%	58.05
2017	16.16%	41.13	14.06%	54.82
2018	11.49%	39.49	12.17%	54.31
2019	8.99%	39.53	12.36%	54.24
2020	8.06%	40.32	10.97%	46.42
2021	7.64%	41.66	10.59%	28.35
2022	7.58%	43.36	9.43%	27.41

The table presents both the percentage of missing values and the mean of existing values for each year in the sample.

Appendix 2: **Average ESG score**



The figure presents the average ESG score of the different providers. The grey line represents Sustainalytics, while the blue line represents Refinitiv.

Robustness and model tests

Breusch-Pagan Test: Heteroscedasticity

Breusch-Pagan Test is to check our sample for homoskedasticity. If heteroskedasticity is present and not controlled for could lead to biased results. Rejection of H_0 would indicate heteroscedasticity is present in our sample, and robust standard errors should be considered.

Appendix 3: **Breusch-Pagan Test**

Model		$\chi^2 - Statistic$	$P - value$	H_0:
ESG Score		5055.039	0.0000	Reject
ESG rank divided into subsamples	***Normal***	2157.486	0.0000	Reject
	Crisis	108.894	0.0000	Reject
ESG rank divided with dummy for stress periods		5083.126	0.0000	Reject

The table presents a Breusch-Pagan test on our models, and in all models, H_0 is rejected, indicating heteroscedasticity is present in the data. All tests use Refinitiv as a reference and are assumed to hold for Sustainalytics as well.

Breusch-Godfrey Test: Autocorrelation

Breusch-Godfrey Test is to check for autocorrelation in the errors of our models. Rejection of H_0 would indicate autocorrelation is present in our sample, and robust standard errors should be considered.

Appendix 4: **Breusch-Godfrey Test**

Model		$\chi^2 - Statistic$	$P - value$	H_0:
ESG Score		1007.503	0.0000	Reject
ESG rank divided into subsamples	***Normal***	968.588	0.0000	Reject
	Crisis	68.5656	0.0000	Reject
ESG rank divided with dummy for stress periods		1005.373	0.0000	Reject

The table illustrates the results from the Breusch-Godfrey Tests, and in all models, H_0 is rejected, indicating heteroscedasticity is present in the data. All tests use Refinitiv as a reference and are assumed to hold for Sustainalytics as well.

Restricted F-Test – Pooled vs. FE

The F-test is conducted to test if any entity/time-specific (μ_i) intercept is non-zero. If H_0 is rejected, we prefer FE over Pooled regression as there will be a significant goodness-to-fit in FE. Results from the table below suggest FE will be a better model for our sample.

Appendix 5: **F-Test**

Model		$F - Statistic$	$P - value$	H_0:
ESG Score		13.341	0.0000	Reject
ESG rank divided into subsamples	***Normal***	12.244	0.0000	Reject
	Crisis	1.755	0.0000	Reject
ESG rank divided with dummy for stress periods		13.292	0.0000	Reject

The table illustrates the F-Tests, and in all models, H_0 is rejected, indicating fixed effects are non-zero. All tests use Refinitiv as a reference and are assumed to hold for Sustainalytics as well.

Hausman test: RE vs FE

The Hausman test checks for unobserved cross-section heterogeneity in the data. If heterogeneity is present, the predictor will be correlated with the error term (endogeneity), e.g., $Cov(X_{it}, u_{it}) \neq 0$. Hence, our estimations might be will inconsistent and biased. When dealing with panel data, the Hausman test will give an indicator that FE or RE is preferred (Brooks, 2019). The H_0suggest RE would be the preferred model, and a rejection of H_0 would indicate FE is desirable.

Appendix 6: **Hausman Test**

Model		$\chi^2 - Statistic$	$P - value$	H_0:
ESG Score		1817.1205	0.0000	Reject
ESG rank divided into subsamples	***Normal***	1006.6898	0.0000	Reject
	Crisis	177.1043	0.0000	Reject
ESG rank divided with dummy for stress periods		1765.2683	0.0000	Reject

The table illustrates the Hausman Tests, and in all models, H_0 is rejected, indicating unobserved entity-specific heterogeneity. All tests use Refinitiv as reference and are assumed to hold for Sustainalytics as well.

Variance Inflation Factor (VIF)

VIF is conducted to test for multicollinearity in our data. The problem occurs when two or more of the independent variables are highly correlated with each other, which might lead to a wrong interpretation of the result as standard errors are too high, resulting in insignificance when there is significance. In the table below, interaction checks for multicollinearity between the interaction terms and ranks and indicates strong multicollinearity between the ranks and interaction between ranks and expense. To limit the possibility of biased results, we omit this interaction term. After controlling for VIF, we obtain close to 1 for all our variables, indicating no serious multicollinearity in our model.

***Appendix 7*: Variance Inflation Test**

<table>
<tr><th rowspan="2">Variable</th><th rowspan="2">Interaction</th><th colspan="2">Model 1</th><th rowspan="2" colspan="2">Model 2</th></tr>
<tr><th>Normal</th><th>Crisis</th></tr>
<tr><td>Low Rank</td><td>11.792</td><td>1.230</td><td>1.280</td><td colspan="2">1.266</td></tr>
<tr><td>High Rank</td><td>9.430</td><td>1.254</td><td>1.267</td><td colspan="2">1.272</td></tr>
<tr><td>Low Rank x M.Ret</td><td>1.200</td><td>1.239</td><td>1.236</td><td colspan="2">1.225</td></tr>
<tr><td>High Rank x M.Ret</td><td>1.163</td><td>1.221</td><td>1.260</td><td colspan="2">1.209</td></tr>
<tr><td>Low Rank x Exp</td><td>12.083</td><td>-</td><td>-</td><td colspan="2">-</td></tr>
<tr><td>High Rank x Exp</td><td>9.270</td><td>-</td><td>-</td><td colspan="2">-</td></tr>
<tr><td>Low Rank x Covid/Conflict</td><td>-</td><td>-</td><td>-</td><td>1.043</td><td>1.041</td></tr>
<tr><td>High Rank x Covid/Conflict</td><td>-</td><td>-</td><td>-</td><td>1.056</td><td>1.047</td></tr>
</table>

The table present VIF test statistics. Control variables are also in the models, but only variables of interest are presented in the table. The results indicate multicollinearity will not be a problem in the models. All tests use Refinitiv as a reference and are assumed to hold for Sustainalytics as well.

Appendix 8: Robustness of result with fund family as a fixed effect

Appendix 8: The tables show the results from the robustness tests where we use fund family as a fixed effect rather than fund fixed effect. Models 1, 2, and 3 represent the same models conducted in the main analysis.

Model 1: Fund family fixed effect

	(1)	(2)	(3)	(4)
Dep. Variable	Net flow	Net flow	Net flow	Net flow
No. Observations	137568	137568	137568	137568
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0595	0.0593	0.0595	0.0595
F-statistic	1084.0	1081.9	867.90	867.92
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000
Monthly Return	0.1226***	0.1151***	0.1000***	0.0997***
	(10.597)	(8.3975)	(10.722)	(10.504)
Morningstar Rating	0.7957***	0.7964***	0.7980***	0.7962***
	(22.603)	(22.413)	(22.421)	(22.351)
Log Size	-0.2553***	-0.2558***	-0.2565***	-0.2555***
	(-18.750)	(-18.802)	(-18.920)	(-18.824)
Expense	-1.0897*	-0.9433	-0.7504	-0.8886
	(-1.6662)	(-1.4937)	(-1.1725)	(-1.4373)
Std Dev	-0.0239***	-0.0226***	-0.0204***	-0.0224***
	(-3.6081)	(-3.3999)	(-3.3059)	(-3.4895)
Log Age	-0.1664***	-0.1694***	-0.1745***	-0.1705***
	(-4.2983)	(-4.3439)	(-4.5159)	(-4.4213)
ESG Score	-0.0998	-0.0633		
	(-1.1818)	(-0.5486)		
ESG * M. Return	-0.0386**	-0.0243		
	(-2.3284)	(-1.0856)		
Low ESG			-0.0796	0.0197
			(-1.6008)	(0.4037)
High ESG			0.0136	-0.0532*
			(0.4284)	(-1.6561)
Low ESG x M. Return			0.0226***	0.0244***
			(3.4408)	(3.0094)
High ESG x M. Return			0.0165**	0.0117*
			(2.4121)	(1.7296)
FE	Yes	Yes	Yes	Yes

Model 2: Fund family fixed effect

	Normal		Crisis	
	(1)	(2)	(3)	(4)
Dep. Variable	Net flow	Net flow	Net flow	Net flow
No. Observations	132221	132221	6422	6422
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0607	0.0607	0.0283	0.0308
F-statistic	852.49	852.20	17.927	19.583
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000
Low ESG	-0.0777	0.0017	-0.0015	0.3818***
	(-1.5082)	(0.0354)	(-0.0258)	(3.1010)
High ESG	0.0137	-0.0410	-0.0454	-0.3365***
	(0.4062)	(-1.2417)	(-0.3133)	(-2.8380)
Monthly Return	0.1017***	0.1013***	0.0549**	0.0475*
	(10.623)	(10.354)	(2.0471)	(1.8618)
Morningstar Rating	0.8059***	0.8045***	0.4926***	0.4838***
	(22.421)	(22.373)	(6.0021)	(6.3534)
Log Size	-0.2594***	-0.2585***	-0.1450***	-0.1450***
	(-18.877)	(-18.830)	(-4.9187)	(-5.1001)
Expense	-0.6681	-0.7926	-3.6002	-3.6012
	(-1.0547)	(-1.2916)	(-0.9600)	(-0.9393)
Std Dev	-0.0197***	-0.0213***	-0.0369***	-0.0432***
	(-3.0612)	(-3.2181)	(-2.8605)	(-3.1242)
Log Age	-0.1851***	-0.1817***	0.1401	0.1462
	(-4.8675)	(-4.7985)	(1.4652)	(1.4933)
Low ESG x 1M Return	0.0213***	0.0240***	0.0248***	0.0456***
	(2.9316)	(2.8794)	(4.1585)	(3.8397)
High ESG x 1M Return	0.0183**	0.0119	0.0149	-0.0003
	(2.1220)	(1.4515)	(1.0680)	(-0.0470)
FE	Yes	Yes	Yes	Yes

Model 3: Fund family fixed effect

	(1)	(2)
Dep. Variable	Net flow	Net flow
No. Observations	137568	137568
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0595	0.0596
F-statistic	620.09	621.04
P-value (F-stat)	0.0000	0.0000
Low ESG	-0.0756	0.0003
	(-1.4499)	(0.0061)
High ESG	0.0142	-0.0384
	(0.4252)	(-1.1886)
Monthly Return	0.1001***	0.0994***
	(10.750)	(10.418)
Morningstar Rating	0.7981***	0.7961***
	(22.400)	(22.337)
Log Size	-0.2565***	-0.2556***
	(-18.902)	(-18.881)
Expense	-0.7494	-0.8788
	(-1.1689)	(-1.4282)
Std Dev	-0.0204***	-0.0223***
	(-3.2948)	(-3.4743)
Log Age	-0.1745***	-0.1709***
	(-4.5175)	(-4.4358)
Low ESG x M. Return	0.0221***	0.0269***
	(3.4423)	(3.2288)
High ESG x M. Return	0.0164**	0.0092
	(2.2020)	(1.3101)
Low ESG x Covid	-0.1405	0.4694**
	(-1.4604)	(2.5042)
High ESG x Covid	-0.1713**	-0.2292*
	(-2.1889)	(-1.9423)
Low ESG x Conflict	-0.0150	0.2277*
	(-0.2088)	(1.7234)
High ESG x Conflict	0.1480	-0.2857***
	(0.9464)	(-4.6774)
FE	Yes	Yes

Appendix 9: Robustness of result with increased ESG ranks

Appendix 9: The tables show the results from the robustness tests where we increase the High and Low rankings criteria to the top and bottom 20%. Models 1, 2, and 3 represent the same models conducted in the main analysis.

Model 1: ESG ranks at 20%

	(1)	(2)	(3)	(4)
Dep. Variable	Net flow	Net flow	Net flow	Net flow
No. Observations	137568	137568	137568	137568
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0601	0.0597	0.0597	0.0597
F-statistic	1090.7	1082.0	866.06	866.23
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000
Monthly Return	0.1162***	0.1082***	0.0963***	0.0963***
	(9.9303)	(7.8515)	(9.9126)	(9.6485)
Morningstar Rating	0.9236***	0.9261***	0.9266***	0.9258***
	(25.152)	(25.027)	(25.133)	(25.042)
Log Size	-0.6833***	-0.6914***	-0.6889***	-0.6918***
	(-18.091)	(-18.485)	(-18.392)	(-18.580)
Expense	-3.3594**	-3.4444**	-3.4667**	-3.4248**
	(-2.1388)	(-2.1916)	(-2.1995)	(-2.1799)
Std Dev	-0.0195***	-0.0185**	-0.0183**	-0.0185**
	(-2.6221)	(-2.5043)	(-2.5052)	(-2.5030)
Log Age	-1.1192***	-1.0957***	-1.1054***	-1.1019***
	(-5.4853)	(-5.3478)	(-5.4161)	(-5.3898)
ESG Score	-1.2523***	-0.1416		
	(-4.1186)	(-0.8670)		
ESG * M. Return	-0.0375**	-0.0213		
	(-2.4185)	(-1.0059)		
Low ESG			0.0547	0.0650*
			(1.2462)	(1.6560)
High ESG			-0.0718*	0.0133
			(-1.7874)	(0.3921)
Low ESG x M. Return			0.0083	0.0095
			(1.5094)	(1.4731)
High ESG x M. Return			-4.543e-05	0.0021
			(-0.0079)	(0.4443)
FE	Yes	Yes	Yes	Yes

Model 2: ESG ranks at 20%

	Normal		Crisis	
	(1)	**(2)**	**(3)**	**(4)**
Dep. Variable	Net flow	Net flow	Net flow	Net flow
No. Observations	132221	132221	6422	6422
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0600	0.0600	0.0438	0.0466
F-statistic	836.90	836.31	24.379	26.028
P-value (F-stat)	0.0000	0.0000	0.0000	0.0000
Low ESG	0.0608	0.0437	-0.0852	0.3369***
	(1.3831)	(1.1507)	(-0.2173)	(4.1905)
High ESG	-0.0871**	0.0177	-0.3181*	-0.2512***
	(-2.0807)	(0.5064)	(-1.7534)	(-4.2886)
Monthly Return	0.0969***	0.0975***	0.0372	0.0269
	(9.5486)	(9.2556)	(0.7889)	(0.6278)
Morningstar Rating	0.9298***	0.9293***	0.5751***	0.5445***
	(24.605)	(24.549)	(7.4137)	(7.8574)
Log Size	-0.6915***	-0.6944***	-2.3304***	-2.2970***
	(-17.852)	(-18.038)	(-6.0378)	(-6.0944)
Expense	-3.8142**	-3.7827**	2.9074	3.5050
	(-2.3800)	(-2.3706)	(0.4799)	(0.6001)
Std Dev	-0.0167**	-0.0166**	-0.0525***	-0.0660***
	(-2.3249)	(-2.3001)	(-2.8452)	(-2.9896)
Log Age	-1.0551***	-1.0551***	-2.6575*	-2.2382
	(-5.1109)	(-5.0925)	(-1.7083)	(-1.4449)
Low ESG x M. Return	0.0105*	0.0102	0.0007	0.0221**
	(1.7819)	(1.4793)	(0.0667)	(2.5389)
High ESG x M. Return	0.0034	0.0038	-0.0058	-0.0131**
	(0.5535)	(0.7752)	(-0.4175)	(-2.2343)
FE	Yes	Yes	Yes	Yes

Model 3: ESG ranks at 20%

	(1)	**(2)**
Dep. Variable	Net flow	Net flow
No. Observations	137568	137568
Cov. Est.	Driscoll-Kraay	Driscoll-Kraay
R-squared	0.0598	0.0599
F-statistic	619.04	620.26
P-value (F-stat)	0.0000	0.0000
Low ESG	0.0584	0.0477
	(1.2863)	(1.2084)
High ESG	-0.0757*	0.0254
	(-1.7858)	(0.7241)
Monthly Return	0.0964***	0.0958***
	(9.9486)	(9.5697)
Morningstar Rating	0.9261***	0.9254***
	(25.062)	(25.098)
Log Size	-0.6888***	-0.6918***
	(-18.353)	(-18.512)
Expense	-3.4591**	-3.4200**
	(-2.1878)	(-2.1713)
Std Dev	-0.0184**	-0.0188**
	(-2.4939)	(-2.5073)
Log Age	-1.1023***	-1.1018***
	(-5.4036)	(-5.3872)
Low ESG x M. Return	0.0079	0.0111*
	(1.4705)	(1.7539)
High ESG x M. Return	0.0006	0.0007
	(0.0983)	(0.1399)
Low ESG x Covid	0.1355**	0.4757***
	(2.3729)	(4.6843)
High ESG x Covid	0.0594	0.0223
	(0.3637)	(0.2034)
Low ESG x Conflict	-0.2465***	-0.0078
	(-4.2865)	(-0.0757)
High ESG x Conflict	0.0642	-0.3359***
	(0.5701)	(-7.0574)
FE	Yes	Yes