ESG ETFs: Just a label, or the way forward?

Erasmus University Rotterdam

Financial Economics

J.T.J. de Fluiter Balledux

427051

Supervisor: E. Smajlbegovic

August 6, 2021

Abstract

As the global demand for ESG investments increases, so does the call for a transparent investment process. Currently, the majority of ESG investments is done through ESG ETFs but there is a lot of unclarity about the price investors pay for these products, as measured by the ETF's expense ratio, as well as their quality. This thesis aims to find out to what extent the height of the expense ratio determines the quality of the ESG ETF product, looking at variables such as financial return and ESG performance. The ESG ETFs are also compared to a universe of benchmark ETFs and are assessed in terms of flow response to past performance. The analysis finds no relation between expense ratios and financial return. ESG ETFs are also found to significantly underperform benchmark ETFs in terms of financial return, however, this gap can be mitigated through buying more expensive ESG ETFs. In terms of ESG performance the results are mixed, with different ESG score proxies resulting in different results. The implications are similar however: more funding to the ESG ETF portfolio manager, either through a higher expense ratio or a bigger sized fund, leads to a higher quality ESG product. Lastly, despite an increased effort to make investments not just about financial return the analysis finds that especially ESG ETFs are judged critically based on their financial return.

I. Introduction

As the dangers of global warming and climate change become more imminent by the day, it can be said beyond reasonable doubt that action has to be taken sooner rather than later. The financial industry tries to take part in this action by promoting sustainable investments, and both local and national governments try to subsidize green initiatives as well. As an example, the EU Covid recovery fund and long term budget plans (total value of €1.8 trillion) will see at least 30% being spent on climate concerns (Abnett & Green, 2020). Moreover, all spending done within these programs must in no way harm the EU's goal of net-zero GHG emissions by 2050. These policies are just one out of many examples of the increased attention posed on green investments which are part of a general trend in which investments are not solely judged based on their financial performance but also based on their sustainable performance. Using investment vocabulary, recent investment developments focussing on sustainable performance are trying to adhere to criteria regarding Environmental, Social and Governmental factors, or ESG in short.

Currently one of the most popular sustainable investment vehicles are ESG ETFs, which are listed entities which allow an investor to buy a multitude of different ESG related products (mostly stocks) with one transaction. Over the coming years, ESG ETFs are expected to gain even further in popularity. A report by PWC predicts that by 2025, between €5.5 trillion and €7.6 trillion worth of assets managed by European mutual funds (ETFs included) will be ESG assets, representing 41% and 57% of the total assets under mutual fund management respectively. This is an increase of roughly 2.5 – 4 times the number which was measured in December 2019, when only 15.1% of assets under mutual fund management were recognized as an ESG asset (PWC Research, 2020). Adding ESG related criteria and labels to investments is a positive trend as it increases the overall awareness towards this type of investing. However, it also brings some downsides.

One of the current issues within ESG investments is the concept of greenwashing: The deliberate publication of misinformation about a firm's ESG performance (Delmas & Burbano, 2011). In more simple terms, greenwashing could be a company who claims to be very responsibly in terms of ESG performance, but in reality is not. Within the universe of ESG ETFs, it could be an ETF which claims to be ESG, but whose underlying securities are not. As the market for ESG investments grows, so do the scandals surrounding greenwashing with Thompson (2019) suggesting that the problem might increase even further as managers aim align their company's perceived image with the ever tightening ESG rules and regulations.

By falsely claiming that an investment ranks well according to the ESG criteria, investments are made in the wrong assets, failing to add anything towards the battle against climate change in global warming. In order to further improve the transparency involved with the process of ESG ETF investments, it is beneficial to all parties involved within the industry of ESG ETFs to know how their underlying qualities differ from non-ESG ETFs. When can an ETF be classified as ESG and are there differences in ESG ratings between these ETFs? Moreover, investors perhaps wonder if there is a premium involved for ESG ETFs compared to non-ESG ETFs and whether it is possible to pay up front for a higher quality ESG product?

This thesis will aim to answer these question and will therefore add to the ever-growing literature on (ESG) ETF dynamics. The framework of the analysis will be based on the classic fee return research on mutual funds. After all, the fee of a fund is the price you pay on top of your initial investment, and economic theory suggests that a higher price should lead to some type of compensation (Layard, 1994; Markowitz, 1952). Based on a broad set of academic literature, there is currently a consensus that higher fees will not lead to a higher financial return and is often even a negative predictor of financial return (Mansor, Bhatti & Ariff, 2015; Vidal, Vidal-Garcia, Lean & Uddin, 2015). This thesis will add to the existing literature framework of fee relations by examining to what extent an investor is paying for other types of compensation such as ESG performance. The final part of the analysis will be devoted to the investors' response towards all the variables of interest by assessing how ESG ETF fund flows respond and whether this is different from regular ETFs.

The remainder of the thesis will be structured as follows: Section II will cover the relevant theory which will lead to a set of testable hypothesis. Section III will then describe the data used, after which section IV will cover the method. The results are put in section V, after which section VI will further discuss and interpret the results. Section VII concludes.

II. Theoretical Framework

The following section will provide the academic framework on which the remainder of the research will be built. It will aim to provide a clear cut definition of (ESG) ETFs and their underlying mechanics. Moreover, it should provide the reader with the tools to formulate the first ideas about the possible relationship between ESG ETF fees and the ESG performance of the ETF.

Before going into the relevant economic theories and academic literature, the concept of mutual funds and ETFs will briefly be explained in the first paragraph. The second paragraph

will then use the aforementioned economic theory as well as academic literature to formulate the baseline ideas of ETF fee mechanics and why they can vary between different funds or ETFs. The baseline ideas will be developed upon in section three, where a link with the existing academic theory will be made in order to further formulate the different hypotheses this research will test.

Understanding mutual funds and ETFs

As suggested by Ferri (2009), in order to understand ETFs it is useful to become familiar with the concept out of which they have evolved: the mutual fund. With this in mind the following sub-chapters will cover both concepts. Mutual funds will be briefly introduced first, after which the specific characteristics of close- and open-end mutual funds will be used to give some colour to the concept of Exchange Traded Funds (ETFs).

The reason to also report on mutual fund literature within a research devoted to ETFs is twofold: **1)** Because ETFs bear many similarities to mutual funds and can thus serve as a clarifying first step; **2)** Because the mutual fund literature is currently much more advanced than the ETF literature, despite the recent gain in popularity of ETFs over the last two decades. It should be noted, however, that despite the many similarities between ETFs and mutual funds there are also vast differences, of which the most important ones will be discussed as well.

Mutual funds

The origination of the contemporary mutual fund goes back all the way to the end of the 18th century when in 1774 a Dutch broker invited investors to subscribe to a trust called '*Eendragt Maakt Magt*' (Rouwenhorst, 2004). Investors with a subscription to the trust were able to buy into a limited selection of shares which the fund had issued. The money paid by the investors for the shares was deposited into the trust, which would then invest the capital on behalf of the investors in foreign government bonds and bank and plantation related loans. Shareholders were entitled dividend and could sell their shares to other investors if they wanted to, or they could increase their position by buying shares from other shareholders (Rouwenhorst, 2004). The way in which the trust was operated bears many similarities to what we nowadays refer to as closed-end mutual funds.

Closed-end mutual funds also issue a fixed number of shares which give their owner a stake over a collection of tradable securities (Elton & Gruber, 2013; Rouwenhorst, 2004). Selling or buying a share in a closed end fund is possible as long as the exchange on which the

fund is listed is open. Because they are listed on exchanges, the price dynamics of closed-end funds are based on supply and demand (Elton & Gruber, 2013). One of the consequences of this pricing mechanism is that closed end fund share prices often trade at a premium or discount to the Net Asset Value (NAV) of the underlying portfolio of securities (Boudreaux, 1973).

Besides the closed-end mutual funds there are also open-end mutual funds, which are more prominent than their closed-end counterparts both in quantity and size of the funds (Boudreaux, 1973; Elton & Gruber, 2013). The main difference with closed-end mutual funds is that open-end mutual funds are not restricted to a certain quantity of shares. If an investor wishes to invest in an open-end fund then the fund can simply issue new shares if needed. As written before, close-end mutual funds only have a limited number of shares which can be traded (Elton & Gruber, 2013). Moreover, the share price of an open-end mutual fund is equal to the fund's NAV. Once a day this value is determined and published, after which all the buy and sell orders of that day will be processed simultaneously. A similarity between closed-end and open-end mutual funds is that both varieties will grant the investors access towards a basket of securities through a single transaction.

Mutual funds are often run by asset managers or other types of financial institutions such as investment banks, wealth managers or pension funds. These companies collect the money transferred by the investors, manage the holdings of the fund and pay dividends if the fund holds any dividend paying stocks (Sialm & Starks, 2012).

Exchange Traded Funds

Exchange Traded Funds, abbreviated as ETFs, came into existence around 1993 (Ferri, 2009). Although many different varieties of ETFs are currently being traded, they all share a common principle: they are a low-cost and liquid investment entity which represents a basket of multiple securities (Deville, 2008). The first ETFs simply tracked traditional equity indices such as the S&P500 or the Russell 2000 and were created to offer investors a highly diversified, liquid and low-cost possibility to invest into the entire index at once. Nowadays, ETFs come in many flavours and varieties, covering not only equities but also currencies, credits and commodities (Hill, Nadig & Kougan, 2015). Currently, ETF providers are also not simply limiting themselves to the replication of indices anymore, but they have also issued ETFs which represent certain investment themes or industries such as 'Emerging Markets', 'Financials', 'eSports', 'Cannabis' or 'ESG' (Ben-David, Fanzoni & Moussawi, 2021; Deville, 2008; Hill et al., 2015). As noted in the introduction, the scope of this thesis will focus specifically on

ETF which are specialised in ESG investments. The concept of ESG ETFs will further be elaborated upon in a later section.

A paper by Ben-David et al (2021) refers to these types of ETFs as specialized ETFs, whereas index mimicking ETFs are known as broad-based ETFs. Both types of ETFs share many similar advantages towards investors, as they provide an accessible means to increase one's exposure towards a certain subsector or market without having to buy all the different securities one by one.

ETFs are listed entities, just like a closed-end mutual fund, which implies that their shares can be bought or sold as long as the bourse on which they are listed is open (Deville, 2008). The value of the ETF is theoretically based on market supply and demand, but in reality is mainly dependent on the market price of the underlying securities which the ETF holds. Following the theory of market efficiency, one would assume that the market deduced price of an ETF is simply the weighted (depending on the allocation within the ETF's portfolio) average of the underlying securities. If this were not the case then arbitrage opportunities would persist, resulting in a premium or discount compared to the ETF's NAV.

Engle & Sarkar (2006) find that this is indeed the case. ETF prices deviate very seldomly from the fair value of the portfolio of underlying securities. The authors find that there is a slight premium / discount compared to the fair value of the underlying securities for international ETFs, which they explain by referring to the limited liquidity of these ETFs. Generally speaking, however, the value of an ETF is equal to the weighted average of its underlying securities (Engle & Sarkar, 2006). Following these characteristics, ETFs seem at first hand quite similar to close-end mutual funds, but in practice they are a more liquid variety with lower costs whose share price is based on the NAV of the underlying securities and not on market supply and demand of the ETF itself.

Ever since the first ETF launched in 1993, investors' interest towards ETFs has risen exponentially, mainly driven by increasing accessibility towards both retail and institutional investors and the aforementioned benefits of diversity and low costs (Hill et al., 2015). At the end of 1993, the first (and only) ETF held assets which had a total value of $464 million (Ferri, 2009). According to a report written by the Financial Times, the global assets held by ETFs passed the value of $7 trillion in 2020, representing a CAGR of 76.1% between 1993 and 2020 (Flood, 2020). Given the switch from active to passive investing and the aforementioned benefits of ETFs (low costs, accessibility, liquidity, diversity), it is expected that the flows

towards ETFs will increase even further the coming years as the supply of ETF varieties will continue to expand as well.

The dynamics of fund fees

Whenever an investor invests money through a mutual fund or ETF, there are certain costs (fees) that the investor pays on top of the money it wishes to invest. The fee is paid to the institution managing the fund / ETF as a means to cover their operational costs. As a result, the height of the fee might differ depending on the governance structure of the fund (Tufano & Sevick, 1997), but might also be dependent on various research related expenses made by the fund managers (Dellva & Olson, 1998). The following segments will look into mutual fund fees as well as ETF fees and the differences between them. Correspondingly, economic theory and academic literature will be used in order to try to understand why some ETFs charge higher fees than others and what implications this could have on certain performance measures.

The terms 'fee' and 'expense ratio' will both be used throughout the following paragraphs. To clarify, fees are costs that are attached to specific parts of investing in a mutual fund. The term 'expense ratio' is an overarching concept and will be used in a similar fashion as stated by Droms & Walker (2001): The percentage of a fund's total expenses as a fraction of the total Net Asset Value. Following these definitions, combining all the different fees will lead to a total value of expenses, based on which the expense ratio is calculated.

Mutual fund and ETF fees

Mutual funds can charge various types of fees, the exact implementation of which is dependent on the mutual fund manager. Typically, fees are charged both when someone invests into the fund (front-end load charges), or when an investments is withdrawn from the fund (back-end load / redemption charges). On top of this, mutual funds can charge fees that cover their distribution and marketing costs (12b-1 fees) or fees that only have to be paid if an investor redeems their investment within a certain time period (deferred sales charges), the latter of which is used to incentive longer term investments (Dellva & Olsen, 1998).

Generally speaking, ETF fees are much lower compared to mutual fund fees, because they are not as actively managed as mutual funds (Box, Davis & Fuller, 2020) and because ETFs do not charge 12b-1 fees (Boyte-White, 2021). Another significant part of ETF cost dynamics is the fact that ETF investors trade ETF shares on exchanges, meaning that trading costs are usually incurred by brokers as well. This is vastly different from open-end mutual

funds where investors often interact directly with the fund and not through a broker (Hill et al., 2015). As a result of this, mutual funds feel the need to charge investors as a compensation for the trading costs. ETFs generally only publish the expense ratios because they do no really explicitly distinguish between the different fees as is done for mutual funds.

The gap between ETF fees and mutual fund fees has narrowed in recent years. The main explanation is that many new ETFs choose to "track" a specific industry or a specialised subset of companies, activities which require more research on behalf of the ETF manager and, therefore, require more costs which translates into a higher expense ratio for the investors (Box et al, 2020). In response, a news article written by Loder (2019) reports that investors have no problem with this, as they are willing to pay a higher fee for more specialised ETFs, led by the belief that specialised tracking will yield a higher return.

The economic rationale behind different fee structures

There are multiple ways to assess the rationale or justification of different funds charging different types of fees. Using a simple cost-benefit analysis, an investor would only be willing to pay for a certain fee if it felt like the costs of paying said fee are significantly compensated by the utility he / she receives from investing into the fund. Ceteris paribus, an increase in the fee would mean that the investor would only be willing to still invest into the fund if the expected utility after having paid a higher fee would increase by the same amount as the amount by which the fee has increased (Layard, 1994). This framework bears some resemblance to the risk-return analysis of established portfolio theory. In short, for any given level of risk an investor will always look for the asset with the highest return; given a certain level of return, an investor will always prefer the asset with the lowest risk (Markowitz, 1952). In some way, one can view the fee the funds charge towards the investors in a similar way as the risk of an asset. For a given (financial) return of a fund, a rational investor should always look for the fund that charges the lowest fee. Of course the terms 'fee' and 'risk' are not completely interchangeable in this context, but they both bear the same relation to returns.

Given the Markowitz and the cost-benefit frameworks, we would expect some type of compensation as a result of a higher fee. In some ways, paying a higher fee implies that you are taking more risk as there are always cheaper (less risky) alternatives. Knowing this, an investor is only willing to pay a higher fee (taking a higher risk) if it knows that it will receive some type of compensation, for instance through a higher return or through another type of

compensation. From here we can hypothesize that higher fees should be justified by some type of compensation.

Hypothesis 1: A higher expense ratio should lead to a higher 'compensation' towards investors

Fund fees and financial return

One of the most obvious types of compensation when investing into funds is the financial return an investor can get through investing into the fund. Interestingly, the hypothesis as set out in the previous paragraph does not hold in practice when looking solely at financial return. A paper written by Elton, Gruber & Busse (2011) looks into various ETFs tracking the S&P500. They conclude that rational investors should always pick the ETF which has the lowest fees, but that investor rarely do this in practice. Moreover, a paper by Barahona (2020) reports that high fees are a significant predictor of poor index fund performance. In practice this makes sense as well, because if all index funds are tracking the same stocks, an investor will receive the best return by choosing the option which charges the lowest expense ratio. However, the question remains as to why investors are not more selective when it comes to fees.

The literature on mutual fund fees and financial performance is much broader than the available ETF literature, yet, conclusions are similar across both research fields. Generally speaking, higher mutual fund fees are not associated with significantly higher returns, as found by (Gil-Bazo & Ruiz-Verdu, 2008). The authors define the mutual fund market as one which is both competitive and has shown signs of information asymmetry between fund and investor. Oddly enough, the authors find that worse performing funds are often caught charging fees that are either higher than or equal to fees from their better performing counterparts (Gil-Bazo & Ruiz-Verdu, 2008). In a follow up paper, Gil-Bazo & Ruiz-Verdu (2009) use their earlier work to find that funds' pre-fee performance are negatively related to their fees, which was later confirmed by Mansor et al. (2015) and Vidal et al. (2015). Various possible explanations are presented, such as well performing funds being willing to keep fees low due to competition with other well performing funds. Another reason for the negative relation could be the fact that worse performing funds set high fees in order to attract performance insensitive and uninformed investors, i.e. investors who believe that a higher fee will automatically lead to a higher quality product (Gil-Bazo & Ruiz-Verdu, 2009).

Besides this, fees are also notoriously known cut significantly into high pre fee returns. A paper by Mansor et al. (2015) found that fees can reduce the return from as high as 10%

above the benchmark in a pre-fee setting, to as low as 1.69% below the benchmark after accounting for said fees. Vidal et al (2015) find that the negative relationship between fees and financial performance also holds for funds that do not charge any front-end or redemption fees.

These findings, alongside among other papers in the literature, suggest that mutual fund fees are not a positive, and often a negative, predictor of the mutual fund's pre-fee financial performance. From here one can hypothesize that higher fees will not lead to a significantly higher financial performance.

Hypothesis 2: A higher expense ratio will not lead to a higher financial performance

Other types of return: The concept of ESG investing

Since the academic literature seems to indicate that fees are not associated with financial returns, but economic theory suggest that there should be some type of compensation, there must be other types of compensation. Dellva & Ollson (1998) suggest that higher fees can also be a compensation for higher research costs on behalf of the fund manager, leading to a higher quality product. With this in mind, the scope of this thesis will go beyond the relationship between fees and financial performance and will also look at how fees can influence the quality of certain ETFs. A straightforward subset of ETFs that can be used for this are ESG ETFs, with a higher quality ESG ETF implying a higher score in terms of ESG performance of the ETF. In order to properly research the relationship between ETF expense ratios and ESG performance, the concept of ESG investing will first be explained, together with how this can be incorporated into ETFs. The following section will then look at a possible relationship between expense ratios and the ESG performance of ESG ETFs.

ESG is short for Environmental, Social and Governmental and represents a notion that investors should not just look at financial returns but also at the impact of these investments on environmental, social and governmental matters (Friede, Busch & Bassen, 2015). Although the ESG movement has gained most of its popularity over the last two decades, its origins trace back to the 1970s when researchers first started to look at the relationship between Corporate Social Responsibility and Corporate Financial Performance (Friede et al., 2015). Nowadays, ESG is at the heart of investment processes around the globe and can also be recognized as an increasingly important part of governmental policies. A recent example being the €1.8 trillion EU Recovery Fund, also known as the EU's green recovery package. Many experts belief that the diversion from traditional investing towards ESG investing is here to stay long term.

Therefore, it should not come as a surprise that mutual fund and ETF holdings are expected to follow this trend as well. As written earlier, A report by PWC predicts that the relative size of ESG assets, compared to the total amount of assets under mutual fund (ETFs included) management in Europe, will increase between 2.5 – 4 times between 2019 and 2025. As mentioned before, recent years have seen the ETF product diversify into other themes than just indices (Ben-David et al., 2021; Deville, 2008; Hill et al., 2015). One of the results from this is the inception of ETFs related to ESG.

Fully grasping the concept of ESG, one can define ESG ETFs as ETFs of which the underlying basket of securities consist of companies that take the extra step regarding environmental, social and governmental affairs. Over 2020, Morningstar recognized 392 open-end and Exchange Traded Funds available to U.S. investors alone, a 30% increase compared to 2019 (Hale, 2021). Examples of companies that could be included in this ever growing supply of ESG funds and ETFs are Orsted, an manufacturer of renewable energy sources, Sunrun, a solar panel producer or Beyond Meat, producers of plant based meat substitutes. Over the past couple of years, an ever-growing preference towards passive investing, accompanied by increasing awareness towards investments that score high in terms of ESG scores, has led to a fast growing increase of money inflow towards ESG ETFs. Accelerated by increased saving accounts as a result of the Covid-19 pandemic, Morningstar data reports that inflow into "Sustainable" ETFs has quadrupled over the course of 2020, totalling €33.7 billion (Hale, 2021). This was also the first time that ESG ETF inflows were bigger than inflows into the more traditional open-end ESG funds (€17.4 billion in 2020).

Fund fees and ESG performance

The concept of looking not just at financial performance but also at ESG performance of ETFs can provide a new dimension towards our fee-compensation framework. If a higher fee is indeed justified by allocating additional funding towards research, then one would expect that ESG ETFs with relatively higher expense ratios would also score better in terms of their ESG performance, compared to ESG ETFs that charge lower expense ratios. Generally speaking, we already know that ESG mutual funds charge significantly higher expense ratios than their conventional mutual fund counterparts, implying that there are definitely certain cost attached to the implementation of ESG investment within funds (Kempf & Osten, 2008). A survey among ESG investors conducted by van Duuren, Plantinga & Scholtens (2015) finds that investors also use ESG assets as a way to manage long term risk across their portfolio. In

this scenario, a higher ESG performance would imply lower risk across the portfolio, meaning that this could be viewed as some type of compensation as well. These findings are also strengthened by the aforementioned article of Loder (2019) which concerns specialised ETFs charging higher expense ratios to refine their investment portfolio. The economic theories of risk-return and cost-benefit, as well as existing literature on the concept of ESG investments suggest that a higher expense ratio on an ESG ETF could be compensated not by means of higher financial return, but by means of higher ESG performance. This leads to the formation of the third and final hypothesis:

Hypothesis 3: A higher expense ratio will lead to a higher ESG performance within a (ESG) ETF

Measuring the investor response: A flow analysis

The final variable of interest will be the extent to which investors respond to the aforementioned variables of interest. Will more expensive ETFs be punished in terms of capital inflow? And are there differences in this relationship between ESG ETFs and non-ESG ETFs? By looking at the ETF's flows, the percentage change in AUM after having taken into account the ETF's return, it is possible to assess how investors respond to certain ETF qualities such as returns, size and expense ratio.

Various research has already looked at flow relationships in the past, covering both mutual funds and ETFs. The most well documented flow relationship is between flow and prior returns. Investors chase well performing mutual funds, resulting in higher flows the period after which mutual funds have reported a high return (Jaine & Wu, 2000; Roussanov, Ruan & Wei, Sirri & Tufano, 2005). More recent papers looked at the same relationship within the ETF universe, and found similar results (Ben-David et al, 2021; Clifford, Fulkerson & Jordan, 2014). Other variables which have been found to have a positive effect on fund inflows are size and media coverage (Sirri & Tufano, 2005), while expense ratio is found to have a significant negative effect (Ben-David et al, 2021).

Based on the available literature, it is clear that the return flow relation is one that is both persistent and positive.

Hypothesis 4: ETF cash flows have a significant positive relation with prior returns

To this date, no research has yet looked into the differences in flow characteristics between ESG ETFs and non-ESG ETFs. This thesis aims to do so by looking if ESG ETF flow's respond differently towards prior returns and expense ratio compared to non ESG ETFs. By assuming that ESG ETF investors are more concerned with sustainability scores than non ESG ETF investors, it could be possible that ESG ETF investors are less sensitive to past performers. In other words, the flows measured over certain ESG ETFs will be less strongly influenced by prior returns, because ESG ETF investors don't just look at returns, but also look at other types of compensation such as sustainability scores or performance.

Hypothesis 5: ESG ETFs are less sensitive to past performance compared to non ESG ETFs

In line with the research of Ben-David et al (2021), we also expect higher expense ratios to have a negative effect on ETF flow. All things equal, rational investors will prefer the cheaper option.

Hypothesis 6: A higher expense ratio will have a significant negative effect on flow

III. Data

The following section will describe the various databases which were used and the data that was extracted from them. The sample selection process will firstly explain how the distinction has been made between ESG ETFs and non-ESG ETFs. The section will then cover the various ESG score databases that were utilized, mainly explaining the process through which the databases have computed their ESG ratings. A comprehensive overview of the subset of ESG ETFs, as well as the availability of certain variables of interest can be found in Appendix A.

Sample selection

In order to select a relevant subsample of ETFs, the Morningstar Direct database was used. Morningstar one of the industry leaders in the field of mutual fund and ETF research and provides useful tools which can be used to dissect the global ETF market (Ammann, Bauer, Fischer & Müller, 2019). Within Morningstar, a selection is made to look for US listed equity ETFs with an inception date of 1-1-2005 or later. The reason to start in 2005 is because the first ESG ETFs started appearing during that year (Kanuri, 2020). Only equity holding ETFs will

be used due to the ESG rating database of Thomson Reuters only providing scores on equities. The reason to only select funds that are listed in the US is because of the reliance on the CRSP Survivor-Bias-Free Database for certain fund characteristics such as returns. By only selecting U.S. listed funds, data consistency across the sample will be ensured.

The total sample of US listed equity ETFs with an inception date on 1-1-2005 or later consists of 1473 ETFs, of which 1115 are classified as index funds according to Morningstar. The method as presented by Candelon, Hasse & Lajaunie (2021) will be used in order to filter out the ESG ETFs. They build on the approach by Nofsinger & Varma (2014), which searches for specific ESG related keywords within the fund names. Since the fund name is the most visible marketing aspect of an ETF, the authors hypothesize that anything related to ESG investments is most likely mentioned in the fund name. This thesis follows a similar approach which starts with a selection of keywords as presented on the USSIF[1] website. Additional terms such as "ESG", "social" and "peace" among others are then added in order to create a set of words that also corresponds to the S and G facets of ESG, rather than being too focused on the E(nvironmental) related keywords. Table 1 presents both the original USSIF keywords as well as the added keywords.

Table 1

ESG keyword overview: USSIF words and additional words

USSIF	Additions
Community	Climate
Ethical	Clean
Green	Environmental
Impact	ESG
Mission	Fossil (fuel free)
Responsible	Governmental
Socially	Peace
Sustainable	Social
Values	Sustainability

[1] List of USSIF words available at: https://www.ussif.org/sribasics

The total selection of keywords is cross-referenced against the names of the earlier acquired sample of US listed equity ETFs with an inception date after 1-1-2005. Through this method 86 ETFs are identified as ESG ETFs.

In order to compare the dynamics of the ESG ETFs to the dynamics of non-ESG ETFs, a universe of benchmark and parents ETFs is selected as well. For each of the 86 ESG ETFs, a benchmark or parent ETF is identified based on information found in the ETF's prospectus or the fund manager's website. Appendix B provides an overview of the ESG ETFs and the benchmark or parent index to which they were matched. Some ESG ETFs were not assigned a benchmark or parent given the nature of the ETF's holdings.

For each of the benchmark or parent indices, as many ETFs as possible were selected in order to create a benchmark ETF that is as large as possible. Appendix C provides an overview of the 45 ETFs that compromise the benchmark universe and the benchmark or parent index that the ETFs represent. Appendix D provides an overview of the availability of the data used for the benchmarks ETFs, similar to what Appendix A showed for the ESG ETFs.

Table 2 presents a static overview of summary statistics of the entire sample (covering ESG, non-ESG and the benchmark ETFs), with the most current Morningstar data as at the end of May 2021. What immediately stands out is the seemingly hypothesis rejecting feature of expense ratios being significantly lower for ESG ETFs compared to the rest of the ETFs within the original sample of 1473 ETFs ($t = -2.39$, $p < 0.05$). Following our compensation hypothesis, one would expect the ESG scores for ESG ETFs to be higher compared to the sample average, resulting in higher expense ratios. There are, however, also reasons why the expense ratio of ESG ETFs could be lower compared to the sample average. Looking at fund age, ESG ETFs are also significantly younger compared to the rest of the sample ($t = -4.84$, $p < 0.01$), implying that they are more recently created. As mentioned above, ETFs have become cheaper as time has progressed, which might be a reason why the more recently created (ESG) ETFs are lower in cost compared to the overall sample. Moreover, there could also be other types of specialized ETFs in the broader sample which could increase the average value for expense ratios.

Another interesting feature of the static data is the skewness of fund size data, as well as the seemingly apparent investor preference towards non-ESG ETFs (as measured by size values across the different samples). Compared to just the index ETFs, the ESG ETFs are found

to be significantly younger (t = -6.95, p<0.01). Compared to the benchmark universe, ESG ETFs are on average found to be significantly lower for size (t = -3.51, p < 0.01) and age (t = -9.28, p < 0.01) while being significantly more expensive in terms of expense ratio (t = 3.46, p < 0.01). The discrepancy in terms of expense ratio would confirm our compensation hypothesis, with investors paying more for a specialized product. Further analysis will have to show whether this also yields a higher financial return, as suggested by Loder (2019), and whether the higher expense ratios also affect the sustainability performance of the ETF. Perhaps unsurprisingly, the ESG ETFs score better across all sustainability measures.

Table 2

Static overview of ETF sample

Mean values	**Entire sample**	**ESG ETFs**	t-stat vs. entire sample
N	1,387	86	
Size in $m	1,659	853	-0.76
expense ratio	.51%	.41%	-2.39**
Fund age	6.44	3.86	-4.84***
Median values	**Entire sample**	**ESG ETFs**	Chi square median test vs. entire sample
N	1,387	86	
Size in $m	126	119	0.047
expense ratio	.50%	.40%	9.56***
Fund age	5.50	2.8	47.02***
Mean values	**Index ETFs**	**ESG ETFs**	t-stat vs. index
N	1,115	86	
Size in $m	1,882	853	*-0.99*
expense ratio	.47%	.41%	*-1.33*
Fund age	7.49	3.86	*-6.95***
Median values	**Index ETFs**	**ESG ETFs**	Chi square median test vs. index
N	1,115	86	
Size in $m	187	119	4.98**
expense ratio	.46%	.40%	3.06*
Fund age	6.58	2.8	57.79***

Mean values	**Benchmark ETFs**	**ESG ETFs**	t-stat vs. benchmark
N	40	86	
Size in $m	29,142	853	*-3.51****
expense ratio	.26%	.41%	*3.46****
Fund age	12.7	3.86	*-9.28****
TR	50.13	53.51	*12.36****
MS1	46.44	47.74	*5.77****
MS2	24.95	24.28	*-3.67****
Median values	**Benchmark ETFs**	**ESG ETFs**	Chi square median test vs. benchmark
N	40	86	
Size in $m	3,620	119	37.51***
expense ratio	.21%	.40%	8.65***
Fund age	11.3	2.8	38.94***
TR	52.20	54.17	55.41***
MS1	46.81	48.32	31.28***
MS2	24.31	23.24	77.96***

Note: T-statistics are positive or negative dependent on the ESG ETFs' scores being bigger or smaller. Chi square test scores are always positive, a higher score indicates a bigger discrepancy between both median values. Expense ratios are as listed in the ETF's prospectus. Fund age is based on an age calculation made on 7th July 2021. All data used except the ESG scores (TR, MS1, MS2) is as was presented in a snapchat overview in Morningstar Direct at the end of May 2021. The ESG scores are as used in the panel data analysis.

** p <0.1 ** p <0.05 *** p < 0.01*

ETF time series data

Within the sample of the 86 selected ESG ETFs and the 45 benchmark ETFs that were listed in the US, various time series data was downloaded through the CRSP Survivor-Bias-Free US Mutual Fund database. Expense ratio information was downloaded for every year in which the selected ETFs were active. Because CRSP has not been fully updated for 2020 data yet, the database was often unable to disclose expense ratio values for 2020. These values were acquired either through Bloomberg or through the ETF's website. The downloaded expense ratios were provided on a yearly basis, and in the final dataset they have been extrapolated to monthly data in order to match the other variables of interest. Monthly returns and AUM figures were downloaded through CRSP as well, together with the Fama-French and momentum factors from the Kenneth French data library. Return and AUM data was available

for 77 out of the 86 ESG ETFs and 26 out of the 45 benchmark ETFs. For the ESG ETF this was due to some ETFs having only been recently added to the CRSP Mutual Fund database, and for the benchmarks ETFs this can be attributed to the fact that not all of them are listed in the US and CRSP is a US listed only platform.

Morningstar Sustainability Ratings

ESG ETF sustainability scores were assembled through two different ways, the first of which is through the Morningstar Direct database. Morningstar provides, among other things, its own Morningstar Sustainability Rating for each ETF. The validity of the score is among the leaders within the fund rating industry and is based on risk scores as published by Sustainalytics. Sustainalytics' ESG risk scores assess to what extent a company's business operations are at risk due to ESG related issues that might arise. For each month, Morningstar calculates the weighted average of ESG risk score based on the underlying holdings of the ETF. The calculated score is somewhere between 0 – 100 with a lower score implying lower ESG related risks and a higher performance in terms of ESG criteria. In practice, most scores are somewhere between 0 – 50. Officially any score below 20 is classified as "low-risk", whereas any score that is 30 or higher can be classified as "high-risk".

The method of calculating fund level ESG scores as just described has been implemented into the Morningstar system from September 2019 onwards. A different calculation was used before that, which yielded significantly different scores. The implications of the old scores are also different (lower scores imply a worse ESG performance), meaning that the two scores can in no way be compared between themselves. Because of this, the Morningstar Sustainability Rating has been split into two different subsets of scores: MS1 and MS2. MS1 covers all ratings before from January 2015 until September 2019, MS2 covers all ratings from September 2019 until December 2020. Morningstar Sustainability data was available for 77 out of 86 ESG ETFs and 23 out of 45 Benchmark ETFs, due to some of the funds not meeting the threshold of having at least 67% of assets under management within the Sustainalytics database, for instance due to international holdings, or due to the fund having only been active for a few months.

Datastream ESG ratings

As an additional means of assessing ESG performance, this research will also manually calculate ESG scores on an ETF level, the exact method of which will be explained in the next

section. The means of calculating is similar to how Morningstar calculates their Sustainability Score, basing the performance on a fund level on the ESG scores of the underlying securities.

For ESG scores on a stock level, Thomson Reuters ESG scores (formerly known as ASSET4 ESG scores) were downloaded. Thomson Reuters calculates the ESG scores based on 400 ESG variables it processes for each company individually, which are then compiled into 178 Thomson Reuters ESG variables. Those variables are assembled into 10 categories, each of which are placed into either the E, S or G pillar. The scores within the pillars depend on how well industry peers rank within the same category, as well as on how many industry peers also ranked within the database. Table 3 gives a full breakdown of the variables used within the Thomson Reuters ESG framework.

Table 3

Thomson Reuters ESG score variable breakdown

Pillar	**Category**	**# of variables**	**Overall weight**
Environmental	Resource use	19	11%
	Emissions	22	12%
	Innovation	20	11%
Social	Workforce	29	16%
	Human rights	8	4.5%
	Community	14	8%
	Product responsibility	12	7%
Governance	Management	34	19%
	Shareholders	12	7%
	CSR Strategy	8	4.5%
Total		**178**	**100%**

Thomson Reuters also takes controversial events into account. The controversy score, combined with the single E, S and G score creates a combined ESG score, which is the value that was used throughout the manual calculation of ETF level ESG scores. The implication of the Thomson Reuters ESG scores works similar to Morningstar's MS1 scores, with higher values implying better ESG performance. Just like with Morningstar ratings, the Thomson Reuters ESG scores are between 0 – 100.

Sustainability ratings over time

Figure 1 visually represents how the average values of all three sustainability measures (TR, MS1, MS2) have developed over time for our entire ETF universe (ESG ETFs and their benchmark ETFs). In terms of longevity of the data, the most informative pattern is given by TR. Ranging from 2010 until the midpoint of 2019, a clear upward trend is noticeable over time, implying that the ETFs for which TR sustainability ratings could be computed have increased their sustainability performance as the years have progressed. The Morningstar ratings show a conflicting story. The MS1 rating gradually declines over time, implying an increasingly worsening sustainability performance for the funds for which this data was available; the MS2 rating declined as well, but here this implies an increasingly improving sustainability performance.

Figure 1

Sustainability measure development over time – entire sample

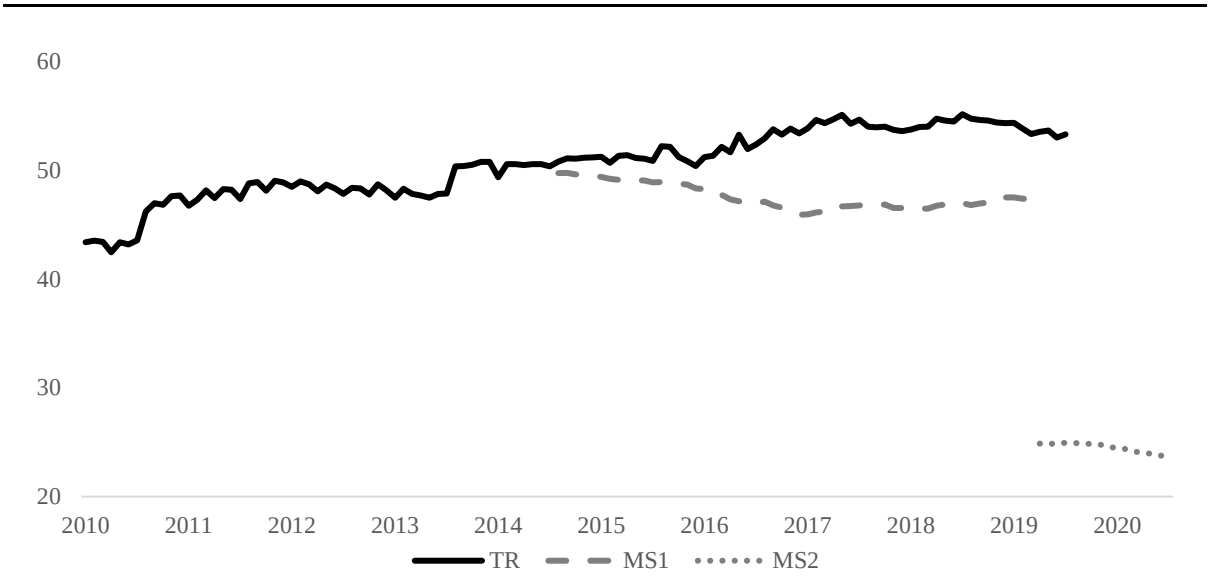
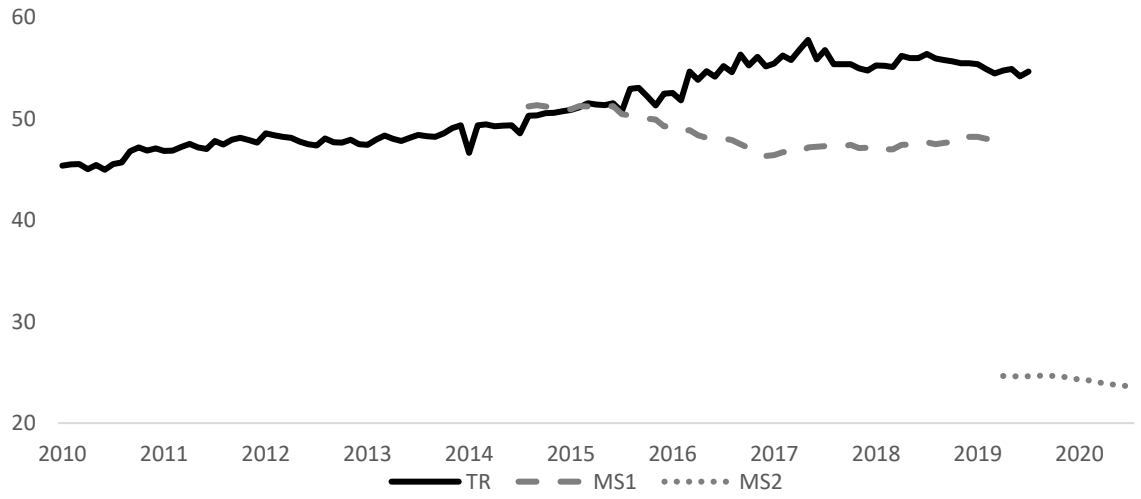


Figure 2 recreates the figure shown in figure 1, but then only showing the ESG ETF subsample. The ESG subset shows more variation within the different sustainability scores. Interestingly, the decrease of the MS1 variable seems to be the strongest within this subsample. Overall, the difference in trends between TR and MS1 once again also proves why different data sources are needed when looking at ESG investment behaviour. At this point in time the

different suppliers are not in line with each other yet, which means that it is necessary to remain critical of the ESG score supplier one is using and to also consider alternative sources.

Figure 2

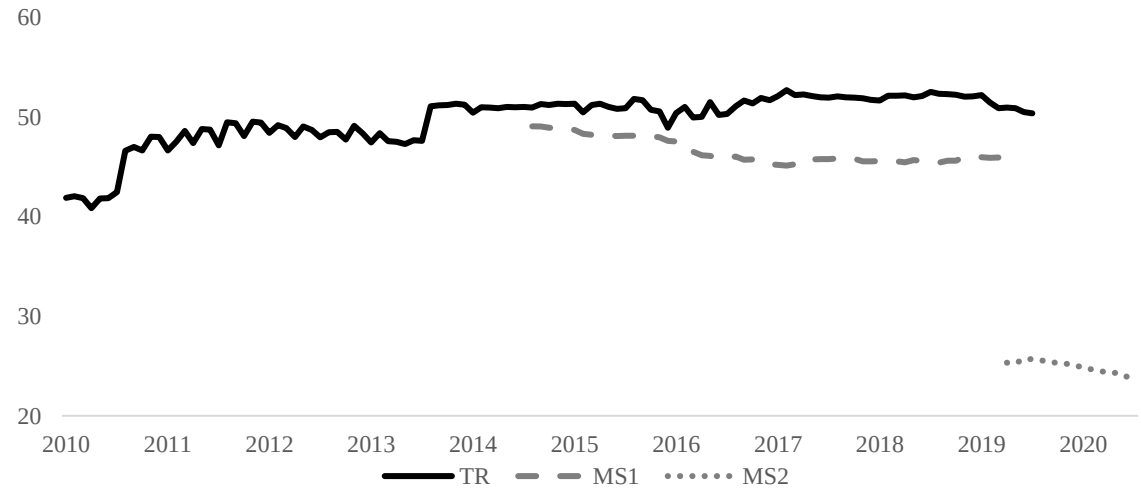
Sustainability measure development over time – ESG sample



Contrary to the ESG sample, the benchmark sample sustainability scores move mostly in line with the combined sample, as is shown in figure 3.

Figure 3

Sustainability measure development over time – Benchmark sample



IV. Method

Financial performance

This thesis aims to follow the method as used in the fee research by Vidal et al (2015) and Gil-Bazo & Ruiz-Verdu (2009), who measured financial performance as the pre-fee abnormal returns. The regression analysis formed to assess the influence of the expense ratio on financial performance can then be described as follows:

$$\alpha_{it} = \delta_{0t} + \delta_1 f_{1t} + \varepsilon_{it} \quad (1)$$

Where α_{it} is the fund's financial performance measured as the pre-fee abnormal return and f_{1t} captures the fund's expense ratio. The size and direction of the financial performance effect is captured through coefficient δ_1. A significant value of $\delta_1 > 1$ implies that an increase in expense ratio has a significantly positive effect on the fund's abnormal returns, confirming that investors pay for higher financial returns for ESG ETFs. Alternatively, a significant value of $\delta_1 < 1$ would mean that investors are undercompensated for higher prices, rejecting the compensation hypothesis. Any insignificant value for δ_1 would imply no relation between expense ratios and financial performance.

In order to also include the benchmark ETFs within the analysis and to account for size effects, because the ESG ETFs differ significantly in size from the benchmark ETFs, an expanded regression will also be run:

$$\alpha_{it} = \delta_{0t} + \delta_1 f_{1t} + \delta_2 \text{AUM}_{2t} + \delta_3 \text{benchmark}_{3t} + \varepsilon_{it} \quad (2)$$

As with (1), α_{it} is the fund's financial performance measured as the pre-fee abnormal return and f_{1t} captures the fund's expense ratio. AUM_{2t} captures the size effect and Benchmark_{3t} is a dummy variable which is 1 when it concerns one of the 45 benchmark ETFs and is 0 when it concerns one of the 86 ESG ETFs. A significant value of $\delta_2 > 1$ implies that an increase in ETF size has a positive significant effect on the fund's abnormal returns, implying that there is some type of return to scale. Significant values of $\delta_2 < 1$ or an insignificant coefficient would confirm that size does not positively influence an ETF's financial return. Regarding the benchmark analysis, any positive significant value for δ_3 would mean that our universe of ETF benchmarks yields higher abnormal returns compared to our ESG universe. The conclusion can also again fall the other way: a negative significant value for δ_3 would

indicate that the universe of ESG ETFs earns a higher abnormal return than their universe of benchmark ETFs. The third and final option, no clear difference in abnormal return between the two universes, will be the case for any insignificant value for $\delta_{3.}$

Abnormal return estimation

Pre-fee abnormal returns will be estimated in two ways: Through an asset pricing model and through a benchmark comparison. The asset pricing model used in this analysis will be Carhart's (1997) four factor model. Factor models such as Carhart's aim to explain stock price variance through exposure towards certain risk factors. The abnormal return of a fund is the difference between the fund's actual return and its predicted return as computed through the Carhart model. The equation of the Carhart model can be noted down as follows:

$$R_{it} - r_{ft} = \alpha_{it} + \beta_{1,it}\, RMRF + \beta_{2,it}\, SMB + \beta 3_{,it}\, HML + \beta_{4,it}\, WML + \varepsilon_{it} \qquad (3)$$

Where $R_{it} - r_{ft}$ is a security's return in excess of the risk free rate, RMRF is the market return in excess of the risk free rate for a given geographic market, and SMB, HML and WML are value weighted investments portfolios that capture the size, value and momentum effect respectively (Carhart, 1997).

Because the ETFs within the sample had different geographic areas in which their holdings were located, multiple variations of the Carhart model were used. More specifically, based on the countries in which the ETF's holdings were located, the ETF was assigned to one of the following 4 categories: United States, Developed world, Emerging Markets, Asia Pacific (Excluding Japan). The fund's abnormal returns would then be calculated with the specific factors that were calculated within that certain geographic category, as is presented within the Kenneth French data library. As a robustness check, Fama and French (1992) abnormal returns were calculated and analysed as well. The Fama and French three factor model is equal to the Carhart model, but then without the momentum factor.

As a second measure of abnormal return, a calculation was made based on how high the ESG ETF's return is compared to their benchmark or parent index, the definition of which is described in the data section. For each return observation R_{it}, where *i* is an ESG ETF indicator and *t* a specific date, the ESG ETF's return is subtracted by the average return of its package of benchmark ETFs that also had a return during that specific time period *t*. The ensuing value will be classified as the ESG ETF's return vs. his benchmark, and is an indication of how the ESG ETF's financial return compared to its selection of benchmark ETFs. Adding this financial

measure will allow this thesis to estimate to what extent the ESG ETFs are an improvement over their benchmark or parent indices on a financial level as well. Moreover, by looking at how expense ratios affect this measure one can tell to what extent ETF investors can pay in order to optimize the return of their ESG ETF investment compared to a possible benchmark ETF.

ESG performance

After analysing how the financial performance our universe of ETFs can be explained, and if there are differences between the ESG ETFs and their benchmark ETFs, the scope of the thesis will move onto the second type of compensation: ESG performance. At first, the effect of expense ratios on ESG performance will be measured through a regression similar to the one used in the financial performance analysis:

$$\varsigma_{it} = \delta_{0t} + \delta_1 f_{1t} + \varepsilon_{it} \quad (4)$$

The main difference with this regression compared to (1) is that this time ς_{it} is the variable of interest, which is a proxy for ESG performance. Two ESG performance proxies (MS1 and MS2) were already introduced in the data section. The third proxy, TR, was manually calculated based on Thomson Reuters ESG scores. The next paragraph will thoroughly explain how the TR score was calculated and how it should be interpreted.

Due to the varying nature of the different ESG proxies, the implication of the δ_1 coefficient differs per measure: a significant value of $\delta_1 > 1$ implies that investors are paying for significantly better ESG performance when looking at the TR and MS1 measures (for MS2 it would imply a worse ESG performance); a significant value of $\delta_1 < 1$ implies the reverse: Investors are undercompensated in terms of ESG performance keeping in mind the price they are paying when looking at the TR and MS1 measure, and overcompensated when looking at the MS2 measure.

In similar fashion to the financial performance analysis, there will also be an extended regression analysis where the benchmark universe and where the size effect that arises as a result of adding the benchmark universe is accounted for:

$$\varsigma_{it} = \delta_{0t} + \delta_1 f_{1t} + \delta_2 \text{AUM}_{2t} + \delta_3 \text{benchmark}_{3t} + \varepsilon_{it} \quad (5)$$

The same intuition applies as at (2), but this within the context of the right hand side variables trying to explain the different sustainability measures that are attached to each of our ETFs. Important to keep in mind here again is that any positive significant coefficient will imply an increase in ESG performance when looking at the TR and MS1 variables, but a decrease in ESG performance when analysing the MS2 variable. If the results are consistent throughout all three methods then the coefficient signs of the MS2 analysis should be the exact opposite of the results of the TR and MS1 analysis.

ESG score proxy

As indicated in the previous section, the value of ς_{it}, the proxy for ESG performance, can be based on three different scores: MS1, MS2 and a proprietary calculation of ESG performance based on Thomson Reuters ESG scores of the underlying securities, to which we will refer as TR. Information on how the MS1 and MS2 sustainability scores were acquired through Morningstar was explained in the data section.

The proprietary TR ESG performance measure was calculated as an alternative to the Morningstar scores, and was done by using holding information from the CRSP Mutual Fund Holding database. The database provides detailed ETF holding information for all US listed ETFs on a monthly basis, which allows us to manually calculate ESG score for the subset of ETFs that are listed in the US.

The available holding data starts in June 2010 and runs until the end of 2019, because both CRSP's holding data and Thomson Reuters' ESG scores are not updated yet properly for 2020. Consequently, some of the more recently created ETFs are omitted from this part of the analysis. In the end, manually calculated CRSP scores are acquired for 50 out of the 86 ESG ETFs within the sample, and 23 out of 45 benchmark ETFs. ESG scores on an ETF levels were calculated as the weighted average of the Thomson Reuters ESG scores of the underlying securities per ETF. If any of the stocks did not have an ESG score available then that stock was not used within the weighted average formula. Similar to how Morningstar approaches the calculation of their sustainability ratings, all observations in which less than 67% of the holdings had an ESG rating were not included in the final sample.

After all data had been downloaded and calculated, a final dataset check was made. Every monthly observation that did not have any information on either financial performance or ESG performance was omitted from the dataset because they did not possess any additional value towards the statistical analysis.

Flow analysis

As hypothesized earlier, one can expect to find a response from investors in the extent to which they allocated capital towards or from ETFs based on the ETF's past financial performance. On top of this, it is of interest to find out whether the certain flow return relations work differently for ESG ETFs compared to their non ESG ETF counterparts. In other words, do ESG investors care as much about financial performance compared to a "regular" ETF investor. In order to properly answer these questions, we start by testing the standard flow return hypothesis.

$$Flow_{it} = \delta_1 l.R_{1t} + \delta_2 l.R_{2t} + \delta_3 l.R_{3t} + \varepsilon_{it} \quad (6)$$

The measure of flow is computed as $100* ((AUM_t – (AUM_{t-1} * return_t)/AUM_{t-1}$, and the three independent variables are the raw ETF returns lagged by one, two or three periods respectively. The reason to include multiple lagged period is because investors do not always respond to past performance immediately, meaning that it might take more than one period for the flow effect to occur. Any significant positive value for our three coefficients will confirm the flow past performance hypothesis. Any significant negative value or nonsignificant coefficient will reject the hypothesis.

Regression (6) will be run two times: Once with the entire sample of ESG and benchmark ETFs, and once with only the ESG universe. The differences in outcome between the two regressions will help to answer to what extent ESG ETFs investors respond differently towards financial performance compared to the broader sample in which the benchmark ETFs are included. The analysis is then expanded to investigate potential exogenous effects regarding expense ratios, fund size and the ETF being an ESG ETF or not.

$$FLOW_{it} = \delta_1 l.R_{1t} + \delta_2 l.R_{2t} + \delta_3 l.R_{3t} + \delta_4 f_{4t} + \delta_5 AUM_{5t} + \delta_6 benchmark_{6t} + \varepsilon_{it} \quad (7)$$

As with previous regressions. AUM represents the ETF's total Assets Under Management, and Benchmark is a dummy variable which is 1 when it concerns one of the 45 benchmark ETFs and is 0 when it concerns one of the 86 ESG ETFs. Similar to (6), the regression will be run two times: once with the entire sample and once only with the ESG ETFs.

V. Results

The following paragraphs will present the regression results from the main analyses that were presented in the methodology section, covering our (expense ratio) analysis on financial and ESG performance and the way in which investors responded by transferring to or withdrawing capital from the ETFs. The reported results and the ensuing implications and analysis provide intriguing insights into the relationship between expense ratios and various compensation measures such as financial performance and ESG performance. Moreover, it brings an interesting addition to the flow performance relationship that currently has mostly been researched into the field of mutual funds.

Financial performance - Carhart

Looking at financial performance, the results from our analysis can split between the two financial performance measures were used (factor model abnormal returns and return vs. a universe of benchmark ETFs). Based on the available literature, it was not expected that there would be a significant relation between expense ratios and our measures of financial performance. The results partially confirm this, as there is indeed no effect when looking at the Carhart abnormal returns. The measure which looked at financial performance as the difference between the ESG ETF's return and the return of a series of benchmarks does find a positive relationship between expense ratios and abnormal fund returns, which means that the answer to our financial performance hypotheses is not a straightforward one. Table 4 shows the results from the factor model analysis.

Across the entire analysis, hardly any significant values are found. Expense ratios as well as the AUM measure never show any type of significance, implying that abnormal returns in no way influence this financial performance proxy. Looking at the logic behind the compensation hypothesis, we can reject any type of relation between expense ratios paid by investors and the ensuing financial performance, as was also suggested by a vast selection of mutual fund literature. In other words, a price paid as an investor for investing into the ETFs is not met with a higher return. Interestingly enough, the benchmark dummy is positively significant in 3 out of 4 regressions in which it is included (always at $p < 0.05$). Differences in Carhart abnormal return between benchmarks ETFs and ESG ETFs vary between 0.5% and 0.34% on the whole. The only insignificant benchmark dummy is in the analysis where the MS2 measure is included. In that last regression, MS2 is positively significant ($z = 2.89$, $p < 0.05$) instead of the benchmark dummy. From the positive significant coefficients on the

benchmark we can conclude that an ETF being a benchmark ETF is usually more financially profitable compared to an ETF being a ESG ETF.

Table 4

Financial performance – Carhart abnormal returns

	Carhart abnormal returns				
ER	0.65 (0.77)	0.73 (1.57)	0.1 (0.87)	0.62 (1.35)	1.12 (1.59)
AUM		-0.010 (-0.37)	0.0064 (0.80)	0.0096 (1.28)	-0.0060 (-0.41)
TR			0.000018 (0.42)		
MS1				0.00012 (0.95)	
MS2					0.0026** (2.89)
Benchmark		0.0050** (1.99)	0.0042** (2.04)	0.0034** (2.30)	-0.0023 (-0.70)
Fixed effects	Time	Time	Time	Time	Time
Sample	ESG	ESG + benchmarks	ESG + benchmarks	ESG + benchmarks	ESG + benchmarks
N	71	92	71	71	92
Adj. R^2	0.16	0.19	0.11	.093	.090

Note: Table output shows coefficients and z- statistics within brackets. Standard errors were clustered at the ETF level; ER is the expense ratio as acquired through CRSP, AUM represents the ETF's Assets Under Management, TR, MS1 and MS2 represent ETF sustainability scores. Benchmark is a dummy variable which is 1 when it concerns one of the benchmark ETFs and 0 when it concerns one of the ESG ETFs.
** p < 0.1 ** p <0 .05 *** p < 0.01*

This in itself is not surprising, as it is generally believed that index funds are one of the safest and diverse investments in terms of long term investment return, and all the benchmark ETFs are index tracking funds. On top of this, ESG ETFs are not only looking to return a financial compensation towards their investors but also a compensation in the form of a sustainable investment. Knowing this, the financial return is not the only variable of interest for these type of ETFs.

The positive significant coefficient on the MS2 variable implies that a higher MS2 score (meaning a lower ESG performance) has a significantly positive impact on abnormal returns. The implication of this relationship seems odd at first sight, especially given the insignificant coefficients for the other two sustainability measures. To put this result into a more clear context, it is useful to assess the time period in which the MS2 variable was available. Starting in October 2019 and lasting until the end of 2020, this measure covers the entirety of the Covid-related market crash and the ensuing recovery. During this time, the financial markets saw levels of volatility not seen since the financial crisis of 2008-09 and many stocks (as well as equity holding ETFs) experienced tremendous drops in prices. Despite the popular belief that ESG investments are more resilient during financial crashes (Folger-Laronde, Pashang, Feor & El Alfy, 2020), there is a lot of recent evidence that this was not the case during the Covid crash. Demers, Hendrikse, Loos & Lev (2020) suggest that it was not ESG ratings that provided resilience to stocks during the Covid-19 market crash, but that instead investments in intangible assets were key in this. The paper written by Folger et al. (2020) looked into ESG resiliency as well, but then specifically within the universe of ETFs. Their results are quite similar, suggesting that high ESG performance did not grant any financial protection.

The positive significant MS2 coefficient adds to the case of ESG investments not being as financially resilient as investors might currently think, and even suggests that ETFs that scored higher in terms of ESG scores were worse off financially during the crisis.

Table 5 (next page) further specifies the MS2 regression between the ESG and the benchmark subsample. From the regression specification, one can conclude that the "MS2 effect" is only found within the ESG ETF subsample. Put differently, during the financial markets crash at the start of the Covid-19 pandemic, ESG ETFs with a better ESG performance did significantly worse in terms of financial performance compared to ESG ETFs with a lower score. All in all, this of course does not show the full picture of ESG investment resiliency, but it does add to the case that ESG investments might not be as crisis proof as some might currently belief.

As a robustness check, the regressions in Table 4 were also run using abnormal returns calculated using Fama and French's three factor model (Appendix E). The main implications do not change much, but overall the effects are less pronounced. Expense ratios are still insignificant, as well as AUM, TR and MS1 variables. The benchmark dummy is still significantly positive, but only in the regression in which the TR value is included ($z = 2.07$, $p < 0.05$). The regression in which the MS2 measure is included again has a strongly significant

positive coefficient for MS2 (z = 2.99, p < 0.01), and a weakly significant negative coefficient for the benchmark (z = -1.68, p < 0.1).

Table 5

MS2 Carhart regression subsample specifications

	Carhart abnormal returns	
ER	0.83	2.77**
	(1.00)	(2.44)
AUM	0.70	-0.000033
	(0.34)	(-0.00)
MS2	0.0029***	-0.0000081
	(2.88)	(-0.02)
Fixed effects	Time	Time
Sample	ESG	Benchmark
N	71	21
Adj. R^2	.12	.068

Financial performance – Benchmark

By solely looking at the Carhart abnormal return analysis, there is no reason to believe that investors are paying for financial performance, both for ESG ETFs and the combined ESG and benchmark ETF universe. In order to expand the scope of the analysis, a benchmark framework was added which touched upon a different measure of financial performance. Within the benchmark analysis, financial performance was seen as the return an ESG ETF had compared to a universe of benchmark and parent ETFs. The results are presented in Table 6. In this analysis, expense ratios are found to be a positively significant predictor in the first two regressions where the sustainability measures are not added, as shown by their respective z-scores of z = 2.20 (p < 0.05) and z = 2.24 (p < 0.05). In economic terms, an increase of .1% in the expense ratio will mitigate the financial underperformance vs the benchmark universe between 1.25% - 2.12%. Once again, the MS2 measure has a positive significant positive (z = 3.09, p < 0.01) and in the same regression the effect of expense ratios is positively significant as well (z = 2.64, p < 0.01).

Table 6
Financial performance – Return vs. benchmark

	Return vs. benchmark				
ER	1.25** (2.20)	1.32** (2.24)	-0.17 (-0.40)	-0.042 (-0.11)	2.12*** (2.64)
AUM		-1.50 (-1.04)	2.19 (0.82)	3.95 (1.16)	-1.04 (-0.94)
TR			0.00016 (1.28)		
MS1				0.00030 (1.54)	
MS2					0.0027*** (3.09)
Fixed effects	Time	Time	Time	Time	Time
Sample	ESG	ESG	ESG	ESG	ESG
N	30	30	19	19	30
Adj. R^2	0.21	0.21	0.37	0.24	0.18

Note: Table output shows coefficients and z- statistics within brackets. Standard errors were clustered at the ETF level; ER is the expense ratio as acquired through CRSP, AUM represents the ETF's Assets Under Management, TR, MS1 and MS2 represent ETF sustainability scores. Benchmark is a dummy variable which is 1 when it concerns one of the benchmark ETFs and 0 when it concerns one of the ESG ETFs.
** p < 0.1 ** p < 0.05 *** p < 0.01*

Contrary to the Carhart abnormal return analysis, the results in Table 6 suggest that within the benchmark framework there are signs of a positive relationship between expense ratios paid and financial compensation received. The addition of the TR and MS1 sustainability proxies as right hand variables seems to mitigate the effect, but this could also be due to the relatively smaller samples that are used within those variables. As a robustness check, Table 7 reports the regression results of the baseline regression again, but then only within the samples in which there is a TR or a MS1 rating.

Table 7

Return vs. benchmark subsample specifications

	Return vs. benchmark	
ER	-0.41 (-0.92)	-0.35 (-0.68)
Fixed effects	Time	Time
Sample	ESG ETFs with a TR score	ESG ETFs with a MS1 score
N	19	19
Adj. R^2	0.37	0.23

The results on the left confirm that in these specific subsets of the sample there is no significant positive relation between expense ratios and abnormal benchmark returns. Most likely, this can be attributed to the fact that these respective subsamples involve too little ETFs (N = 19) for any statistical significant relation to appear. Interestingly enough, within the MS2 subsample expense ratio is found to be significant again, but there the sample size is bigger again (N = 30). The positive significant coefficient of MS2 has a slightly different interpretation here compared to the Carhart abnormal return analysis but in the end it leads to the same conclusion. In this case, it implies that compared to the universe of benchmark ETFs the return of ESG ETFs was significantly lower for higher levels of MS2 (a worse sustainability performance). All things considered, these numbers also seem to indicate that, as an ETF investor, you were worse off during the Covid crash in terms of financial return if you had invested in ETFs which had a high sustainability rating.

As an investor, some important conclusions can be drawn based on the result of table 4, table 6 and the ensuing robustness checks. On the whole, one should not expect to be financially compensated looking at standalone abnormal return. Moreover, ESG ETFs seem to do worse in terms of financial performance compared to a universe of benchmark ETFs. All things considered this makes sense as benchmark ETFs are often seen as very stable and trustworthy long term investment vehicles, and ESG ETFs are (at least in theory) also devoting time and resources towards selecting assets that score well in terms of ESG performance. However, despite the financial underperformance against the universe of benchmarks, the analysis has shown that ETF investors can limit their underperformance through the more expensive ESG ETFs.

By defining financial performance as the return of an ESG ETF minus the benchmark return, there is a significant effect to be found when looking at expense ratios. This would imply that, ceteris paribus, investors can choose to pay to decrease some of the financial

performance deficit that ESG ETFs have compared to their benchmarks. In short, an ESG ETF investor will not be able to pay for financial compensation in terms of absolute abnormal returns, but is able to pay close the financial return gap a little by proactively choosing for a more expensive ESG ETF.

ESG performance

The ESG performance analysis was done three times, once for each one of the different ESG performance proxies. Up front it was hypothesized that the lack of any type of financial compensation would be offset by another type of compensation in the form of higher ESG scores within the ESG ETF universe. The results in Table 8 fail to portray a consistent picture, but the overall implication across the three sustainability proxies is similar.

Table 8

ESG analysis

	TR		**MS1**		**MS2**	
ER	725.54** (2.18)	704.89* (1.76)	-342.30 (-1.48)	-236.60 (-1.31)	182.78 (1.32)	241.67 (1.62)
AUM		-2.69 (-0.36)		8.25* (1.92)		-6.21** (-2.08)
Benchmark		-1.88 (-0.84)		-3.13** (-2.56)		1.33** (2.09)
Fixed effects	Time	Time	Time	Time	Time	Time
Sample	ESG	ESG + Benchmark	ESG	ESG + Benchmark	ESG	ESG + Benchmark
N	50	73	51	74	77	97
Adj. R^2	0.094	0.11	0.11	0.11	0.15	0.14

Note: Table output presents coefficients and z-statistics within brackets. Standard errors were clustered at the ETF level; TR is the manually calculated ESG score based on Thomson Reuters ESG data, MS1 represents the Morningstar Sustainability rating from January 2015 until September 2019, MS2 represents the Morningstar Sustainability rating after September 2019; ER is the expense ratio as acquired through CRSP, AUM represents the ETF's Assets Under Management, Benchmark is a dummy variable which is 1 when it concerns one of the benchmark ETFs and 0 when it concerns one of the ESG ETFs.

** p < 0.1 ** p < 0.05 *** p < 0.01*

Expense ratios are a significant positive predictor of the TR variable, both within the sample of ESG ETFs and the combined sample. The effect is slightly smaller when size and benchmark control variables are included ($z = 1.76$, $p < 0.1$) compared to the regression in which just the ESG ETF universe was taken ($z = 2.18$, $p < 0.05$). Overall, the coefficients account to an increase in the TR score of either 7.3 or 7 for each .1% increase in expense ratio. The other variables of interest are not significant.

At first sight, a different regression output is reported for the MS1 and MS2 analyses. Before interpreting these results, it is useful to recall that the MS1 and MS2 variable have different interpretations in terms of what counts as a high sustainability score. The MS1 variable increases in number as sustainable performance increases, whereas the MS2 variable decreases in number as the sustainable performance increases.

Contrary to the TR analysis, AUM and the benchmark dummy provide significant explanatory power, when looking at the MS1 and the MS2 analyses. For the MS1 analysis, the size effect is positively significant at $p < 0.1$ ($z = 1.92$), whereas the benchmark effect is negatively significant ($z = -2.56$, $p < 0.01$). The MS2 analysis shows opposite coefficients (implying the same effect, as explained in the method section) with the size coefficient being significantly negative ($z = -2.08$, $p < 0.05$) and the benchmark dummy as a positive predictor ($z = 2.09$, $p < 0.05$). Based on the regression output, a bigger sized fund will score better in terms of sustainability performance, and being a benchmark ETF has a significantly negative consequence on sustainability performance compared to the whole sample, something which also already became clear when looking at the summary statistics in table 2. According to the regression output, being a benchmark ETF will lower the sustainability score between 1.33 and 3.13.

One explanation for the difference in coefficients between the TR and the two Morningstar scores could be the origination of the expense ratio variable, which is a yearly variable turned into monthly observations. Because of this, the variations within the expense ratio might not have been big enough between the different years, significantly reducing any explanatory power the right hand variable might have. This would especially be the case for the MS2 variable, which only covers the final two years in the overall data sample. As explained earlier, the sustainability measures, alongside the return and AUM variables, are monthly variables which means that any effect with the expense ratio will only show a significant result if there is enough variability within the expense ratio variable over the timeframe. Since the TR variable ranges from 2010 and 2019 and quite a lot of funds change their expense ratios during their initial years, the TR variable encompasses almost all expense

ratio changes within the different funds. Due to the data availability of the MS1 and MS2 measure (they only start in 2015 and 2019 respectively) there is barely any room for the expense ratio to have changed compared to the universe of the TR analysis.

Interestingly enough, the implications for all three different regressions are the same. An increase in expense ratio, ceteris paribus, means that the ETF manager will have more funds at his disposal which he can use to deliver a better quality product. Similarly, an increase in size, ceteris paribus, will also lead to increased funds available to the ETF manager. Even though both scenarios increase different variables, the implication is the same: A higher amount of available financial resources for an ETF will lead to a better quality product, as indicated by higher sustainability scores across all the three different measures. Looking at the hypothesis, it not just an increase in expense ratio that will lead to a higher ESG performance. The overall effect is much broader, and going from the expense ratio and AUM variables the conclusion is that so far, an increase in available funds will lead to a higher ESG performance. As an investor this means that if you are looking to increase the ESG performance of your ETF investment you can either choose for a more expensive ETF or one that is bigger, ceteris paribus.

How investors have responded: the flow analysis

As a final measure, this research has assessed the extent to which investors responded in terms of capital in- and outflow, dependent on the earlier used variables of interest. Based on the flow performance hypothesis, it was expected that ETF flows would increase as a response to positive past performance and decrease if the ETF performed worse in terms of past financial return. Moreover, based on the assumption that ESG ETF investors would care less about financial return than non ESG ETF investors it was hypothesized that ESG ETFs would be less sensitive to past performance compared to their non ESG ETF counterparts.

The flow return hypothesis is confirmed for our ESG ETF universe, with both one period lagged returns ($z = 3.14$, $p < 0.01$) and returns lagged by two periods ($z = 1.73$, $p < 0.1$) having significant explanatory power over ETF flows. Put more concretely, for every 1% increase in return over the prior period, the flow increased by 0.63%. Looking at the second regression in which the benchmark ETF universe was used, the flow return hypothesis could not be confirmed.

Table 9

Flow analysis

	Flow		
L.Returns	0.63***	0.028	0.042
	(3.14)	(0.89)	(1.03)
L2.Returns	0.39*	0.011	0.016
	(1.73)	(0.57)	(0.36)
L3.Returns	0.12	0.010	0.025
	(0.69)	(0.34)	(0.67)
ER			-7.46
			(-0.89)
AUM			0.29**
			(2.18)
Benchmark			-0.095***
			(-.340)
Interaction variables			
ESG ETF x L.Returns			0.31***
			(2.99)
ESG ETF x L3.Returns			0.097
			(1.22)
ESG ETF x L3.Returns			0.027
			(0.28)
Fixed effects	Time	Time	Time
Sample	ESG	Benchmark	ESG + benchmark
N	61	23	84
Adj. R^2	.058	.15	.051

*Note: Table output shows coefficients and z- statistics within brackets. Standard errors were clustered at the ETF level; Flow is calculated as 100 * ($(AUM_t * return_t) – AUM_{t-1})/AUM_t$, the return variables are the ETF's pre fee returns lagged by one, two or three time periods, ER is the Expense Ratio as acquired through CRSP, AUM represents the ETF's Assets Under Management; Benchmark is a dummy variable which is 1 when it concerns one of the benchmark ETFs and 0 when it concerns one of the ESG ETFs. For the interaction variables, ESG ETF was equal to a benchmark dummy of 0.*

** p < 0.1 ** p < 0.05 *** p < 0.01*

The final regression combined both universes and included some interaction variables to specifically measure the extent in which ESG ETFs respond to past returns compared to the benchmark universe. On the whole, the flow return hypothesis is rejected but ESG ETFs do seem to be more sensitive to past performance as is suggested by the coefficients on the interaction variable with returns lagged by one period ($z = 2.99$, $p < 0.01$).

Despite different expectations up front, ESG ETF investors are more sensitive to past performance than their non ESG ETF counterparts. One part of the explanation could be that within the benchmark ETF universe, investors are less picky in terms of financial return because benchmarks often track an already existing index, which means that the ETF manager merely replicates a product that already exists. In this environment, a higher or lower return can not be attributed to a poorly selected group of stocks. ESG ETFs, on the contrary, are often a specialized self-chosen ESG subset of such an index or sometimes even an entirely independent group of stocks that are deliberately chosen based on their sustainable qualities. In either scenario, the ESG ETF manager chose the investments himself, which means that he or she is more responsible for the ETF's financial return compared to the non ESG ETF universe. On top of this, the belief that ESG investments perhaps do not yield as much financial return can make investors increasingly more aware of this, thus increasing their sensitivity towards the financial performance.

A breakdown by year of the flow return hypothesis within the ESG ETF universe, as reported in appendix F, shows that the effect has been significant throughout the entire sample but that the two most recent years show the first sign of a continuing trend. If anything this shows that even in the most recent years, ESG ETF investors have seen financial performance as a very important requirement of the quality of a possible ESG ETF investment. Despite the critical stance towards ESG ETF returns, on the whole investors have responded positively in terms of flow compared to the benchmark universe as is indicated by a significant negative flow coefficient ($z = -0.099$, $p < 0.01$) for the benchmark dummy. This goes in line well with the earlier reported increasing inflow of capital towards ESG investments. Lastly, ETF size is found to be a positive predictor of ETF flow ($z = 2.09$, $p < 0.05$), an effect which is only found in the whole sample and not by looking solely at ESG ETFs. Apparently, ETF investors are more prone to put more capital into relatively bigger sized ETFs, suggesting some type of herding behaviour among investors.

VI. Discussion

One the whole, if there is anything that stood out in terms of explanatory value of expense ratios is that it still seems to rather limited. Based on our results, and in in line with the paper of Barahona (2020), ESG ETF investors seem to be best off picking an ESG ETF that is low in terms of expense ratio and large in size. That way, you know that the ESG ETF manager has enough funds to deliver a high quality product (as measured in terms of sustainability scores) but you do not pay the price for this as the investor. Perhaps future research could you look specifically into the effect of amount of funding available to a fund manager (measured as expense ratio * AUM) and the quality of the product. At this point there is uncertainty about the compensation one receives for paying a higher expense ratio, but perhaps new insights can be gained by performing the analysis on a broader level through the inclusion of the size of the fund as well.

Besides the conclusion, or lack of conclusion, that was drawn about expense ratios there is also still an important discussion to be had about ESG scores and their consistency or standardization across different ESG score providers. Despite the fact that the implications of our results were similar across our ESG analysis regression, some caution is needed in the interpretation of the coefficients. Despite taking as many measures as possible to achieve the highest data quality, at this point in time sustainability ratings are not standardized yet. A good example of this can be seen in figure 1, where the TR and MS1 measures should have moved in line under the assumption that they were assessing ESG performance similarly.

Of course, part of the added value as an ESG score provider is having a proprietary method of analysis. However, due to the immaturity of the ESG ratings and the manipulation that is still possible from a company perspective (greenwashing), one should always try look beyond the ESG ratings and realize that the product of ESG scores is far from finalized yet. At this point in time, there are no better alternatives than using the ratings that were used for this analysis. However, one can expect Morningstar, Thomson Reuters and possible other ESG data providers to continuously improve on their ESG product because the financial industry will demand them to do so for the unforeseeable future.

The big picture conclusion from all of this is that data quality and data supply are key. The validity of future research will only increase further over time as the ESG rating agencies standardize their methods (in line with credit ratings). Moreover, being able to add a wider variety of ETFs over a longer timeframe will surely also lead to more pronounced conclusions in the future if one were to replicate or expand onto this analysis.

A final suggestion regarding future work comes from combining our new type of compensation (ESG performance) and the flow analysis that was also part of this method. Traditionally, the flow analysis has mostly looked at past financial performance and other price related variables such as expense ratio. However, as the financial product expands to other types of variables such as ESG performance it would not be unsurprising that investor flows also start responding to these variables.

VII. Conclusion

This goal of this thesis was to add to the transparency within the world of ESG ETF investments, by exploring and researching the relation between (ESG) ETF expense ratios and two performance measures in the form of financial performance and ESG performance. The theoretical framework behind this research scope hypothesized that higher expense ratios, the price paid by investors on top of their investment, should be compensated by some type of return towards investors. Based on (mostly) mutual fund literature, it was expected that financial performance would not be positively influenced by higher expense ratios, but that the overall quality of the ETF would increase resulting in higher sustainability scores. Within the scope of the research there was also a chapter devoted to fund flows. After all, it is important to know how investors respond to different variables of interest, and how these reactions might differ between ESG ETFs and non-ESG ETFs. The implications that will follow from this line of questioning will provide valuable insights towards professionals around the financial industry, especially those focussing on investments in ESG ETFs and related products, making a significant contributing towards increasing the transparency encompassing the ever-growing ESG trend.

Based on the statistical analysis, multiple conclusions can be drawn about ESG ETF investments and how their compensation and flow dynamics differ from non-ESG ETFs. First and foremost, expense ratios were found to have no relation with fund returns. On top of this, the analysis found that ESG ETFs underperformed a universe of benchmark ETFs in terms of abnormal return. The abnormal return gap between ESG ETFs and non-ESG ETFs can be narrowed, however, as higher expense ratios were found to significantly reduce the discrepancy between the two samples. For investors this means that in terms of returns they are better of investing in non-ESG ETFs, but that they can at least mitigate the financial underperformance of ESG ETFs by choosing a more expensive ESG ETF.

Perhaps more interesting for ESG ETF investors is whether can also exert some type of influence on the ESG performance of their investment through the price of the ETF. This part

of the analysis technically returned inconsistent coefficients across the different sustainability measures, but interestingly enough the implications across the three ESG proxies was the same. For the proprietary sustainability score, expense ratios were found to have a significant positive effect on their value. This was not the case for the Morningstar ratings, however, but in those subsets it was fund size which was found to positively affect an ETF's ESG performance. Since both variables (expense ratios and size), ceteris paribus, increase the ETF managers' available resources as they increase in size as well, the conclusion regarding ESG performance can be linked to the funding available to the ETF manager. In both scenarios, more resources led to a better ESG performance. As an investor this leaves you with two choices if you want to increase the ESG performance of your investment: either find a more expensive ETF or a bigger one. Perhaps to be expected but still worth mentioning, the benchmark ETFs underperformed the ESG ETFs in terms of ESG performance.

Despite an increasing focus on the sustainable aspects of investment in general, it seems that investments are still being extremely critically judged based on their financial performance. The results of the flow analysis imply that ESG ETF flows are much more sensitive than the universe of benchmarks. Part of the discrepancy can most likely be explained due to the fact that the benchmark ETFs are following an already existing index and have no direct control over which stocks to include or not. Despite this, the flow performance relation has been consistently documented over time and it is surprising to say the least that this effect is stronger within in asset class whose main focus should be the sustainability of the investment and not the financial return. If anything this suggests that despite the recent efforts to increase awareness towards the sustainable aspect of investments, we are still far away from prioritizing sustainability as an investment requirement.

The research was limited in a variety of ways which can be used to improve the paper for future research purposes. A major point was the dependence on CRSP for Mutual Fund Holding information and returns and size information. Because of this, the panel regression sample was limited to ETFs that were listed in the US. The holdings of those ETFs did hold multiple geographic areas but a very big part of the global ETF universe was emitted from the analysis due to this limitation. Sustainability ratings are currently an unfinished product, as was also explained in the discussion. Because the ESG trend has only been a relatively recent phenomenon the analysis was limited to only roughly the last 10 years. As the years progress and more ESG ETFs are created, the method as used in this thesis will be able to paint a clearer picture of ESG ETF dynamics and how they compare to non ESG ETFs.

If anything the limitations of this research as well as the inconsistency in terms of the results found for the different ESG proxies has shown that despite the ESG trend becoming increasingly more important, it is still a concept that is unfinished and barely as easily to research as for instance a classic mutual fund related research. This should not prevent future research to dive into the topic, on the contrary, it should motivate others to pick up where others have stopped in order to further develop the framework in which we can perform and understand sustainable investments. Only then we will come close to ensuring that our investments are not just making a financial impact, but also a sustainable one.

References

Abnett, K., & Green, M. (2020, July 22). EU makes world's biggest 'green recovery' pledge - but will it hit the mark? Retrieved November 29, 2020, from https://www.reuters.com/article/us-eu-summit-climate-change-idUSKCN24N231

Ammann, M., Bauer, C., Fischer, S., & Müller, P. (2019). The impact of the Morningstar Sustainability Rating on mutual fund flows. European Financial Management, 25(3), 520-553.

Barahona, R. (2020). Why can't investors pick the right index fund?.

Ben-David, I., Franzoni, F., Kim, B., & Moussawi, R. (2021). Competition for Attention in the ETF Space (No. w28369). National Bureau of Economic Research.

Boudreaux, K. J. (1973). Discounts and premiums on closed-end mutual funds: A study in valuation. The Journal of Finance, 28(2), 515-522.

Box, T., Davis, R., & Fuller, K. (2020). The Dynamics of ETF Fees. Financial Analysts Journal, 76(1), 11-18.

Boyte-White, C. (2021). Why are etf fees lower than mutual fund fees? Retrieved May 11, 2021, from https://www.investopedia.com/articles/investing/102915/why-are-etf-fees-lower-mutual-funds.asp

Candelon, B., Hasse, J. B., & Lajaunie, Q. (2021). Asymmetric Information in the SRI Market. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of finance, 52(1), 57-82.

Clifford, C. P., Fulkerson, J. A., & Jordan, B. D. (2014). What drives ETF flows?. Financial Review, 49(3), 619-642.

Dellva, W. L., & Olson, G. T. (1998). The relationship between mutual fund fees and expenses and their effects on performance. Financial Review, 33(1), 85-104.

Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. California management review, 54(1), 64-87.

Demers, E., Hendrikse, J., Joos, P., & Lev, B. (2021). ESG did not immunize stocks during the COVID-19 crisis, but investments in intangible assets did. Journal of Business Finance & Accounting, 48(3-4), 433-462.

Deville, L. (2008). Exchange traded funds: History, trading, and research. In Handbook of financial engineering (pp. 67-98). Springer, Boston, MA.

Droms, W. G., & Walker, D. A. (2001). Persistence of mutual fund operating characteristics: returns, turnover rates, and expense ratios. Applied Financial Economics, 11(4), 457-466.

Elton, E. J., Gruber, M. J., & Busse, J. A. (2011). Are investors rational? Choices among index funds. In Investments And Portfolio Performance (pp. 145-172).

Elton, E. J., & Gruber, M. J. (2013). Mutual funds. In Handbook of the Economics of Finance (Vol. 2, pp. 1011-1061). Elsevier.

Engle, R. F., & Sarkar, D. (2006). Premiums-discounts and exchange traded funds. The Journal of Derivatives, 13(4), 27-45.

Fama, E. F., & French, K. R. (1993. Common risk factors in the returns on stocks and bond. Journal of Financial Economics, 33, 3-56.

Ferri, R. A. (2009). The ETF book: all you need to know about exchange-traded funds. John Wiley & Sons.

Flood, C. (2020, September 09). ETF assets reach $7tn milestone. Retrieved March 27, 2021, from https://www.ft.com/content/e59346a7-b872-4402-8943-2d9bdc979b08

Folger-Laronde, Z., Pashang, S., Feor, L., & ElAlfy, A. (2020). ESG ratings and financial performance of exchange-traded funds during the COVID-19 pandemic. Journal of Sustainable Finance & Investment, 1-7.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210-233.

Gil-Bazo, J., & Ruiz-Verdu, P. (2008). When cheaper is better: Fee determination in the market for equity mutual funds. Journal of Economic Behavior & Organization, 67(3-4), 871-885.

Gil-Bazo, J., & Ruiz-Verdú, P. (2009). The relation between price and performance in the mutual fund industry. The Journal of Finance, 64(5), 2153-2183.

Hale, J. (2021). Sustainable Funds US Landscape report. Retrieved March 27, 2021 from https://www.morningstar.com/lp/sustainable-funds-landscape-report

Hill, J. M., Nadig, D., & Hougan, M. (2015). A comprehensive guide to exchange-traded funds (ETFs). CFA Institute Research Foundation.

Jain, P. C., & Wu, J. S. (2000). Truth in mutual fund advertising: Evidence on future performance and fund flows. The journal of finance, 55(2), 937-958.

Kanuri, S. (2020). Risk and Return Characteristics of Environmental, Social, and Governance (ESG) Equity ETFs. The Journal of Index Investing, 11(2), 66-75.

Kempf, A., & Osthoff, P. (2008). SRI funds: Nomen est omen. Journal of Business Finance & Accounting, 35(9-10), 1276-1294.

Layard, P. R. G. (1994). Cost-benefit analysis. Cambridge University Press.

Loder, A. (2019, February 19). Who needs free? Passive fund prices flatline. Retrieved March 27, 2021, from http://www.wsj.com/articles/who-needs-free-passive-fund-prices-flatline-11550592000

Mansor, F., Bhatti, M. I., & Ariff, M. (2015). New evidence on the impact of fees on mutual fund performance of two types of funds. Journal of International Financial Markets, Institutions and Money, 35, 102-115.

Markowitz, H. (1952). The utility of wealth. Journal of political Economy, 60(2), 151-158.

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

PWC Research. (2020) – ESG report - The growth opportunity of the century. Retrieved May 4, 2021, from https://www.pwc.lu/en/sustainable-finance/esg-report-the-growth-opportunity-of-the-century.html

Roussanov, N. L., Ruan, H., & Wei, Y. M. (2020). Mutual Fund Flows and Performance in (Imperfectly) Rational Markets?. Jacobs Levy Equity Management Center for Quantitative Financial Research Paper.

Rouwenhorst, K. G. (2004). The origins of mutual funds.

Sialm, C., & Starks, L. (2012). Mutual fund tax clienteles. The Journal of Finance, 67(4), 1397-1422.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. The journal of finance, 53(5), 1589-1622.

Thompson, J. (2019, May 13). 'Green' funds can fall short of buyers' expectations. Financial Times. https://www.ft.com/content/644c1ec4-39d9-11e9-9988-28303f70fcff?shareType=nongift.

Tufano, P., & Sevick, M. (1997). Board structure and fee-setting in the US mutual fund industry. Journal of Financial Economics, 46(3), 321-355.

Vidal, M., Vidal-García, J., Lean, H. H., & Uddin, G. S. (2015). The relation between fees and return predictability in the mutual fund industry. Economic Modelling, 47, 260-270.

Appendices

Appendix A: Data availability overview of ESG ETF sample

The list consists of all 86 ESG ETFs that were selected through Morningstar Direct. "TR" is an indication of Thomson Reuters ESG score availability; "MS1" and "MS2" indicate availability of Morningstar Sustainability ratings. "Returns" indicate return availability, and "geography" shows which factors were used to calculate the fund's abnormal returns. Information on expense ratios was available for each ETF.

Name	TR	MS1	MS2	Returns	Geography
Adasina Social Justice All Cp Glbl ETF			x	x	Developed
ALPS Clean Energy ETF	x	x	x	x	United States
American Century Sustainable Equity ETF			x	x	United States
American Conservative Values ETF			x	x	United States
Xtrackers S&P SmallCap 600 ESG ETF					
Xtrackers S&P MidCap 400 ESG ETF					
Change Finance US LgCp FossilFuel Fr ETF	x	x	x	x	United States
Changebridge Capital Sustainable Eq ETF			x	x	United States
ClearBridge Dividend Strategy ESG ETF	x	x	x	x	United States
ClearBridge Large Cap Growth ESG ETF	x	x	x	x	United States
Columbia Sustainable Intl Eq Inc ETF	x	x	x	x	Developed
Columbia Sustainable US Equity Inc ETF	x	x	x	x	United States
Direxion MSCI USA ESG Ldrs vs Lggds ETF			x	x	United States
Ecofin Global Water ESG	x	x	x	x	Developed
Etho Climate Leadership US ETF	x	x	x	x	United States
First Trust EIP Carbon Impact ETF	x	x	x	x	Developed
First Trust NASDAQ® ABA Community Bk ETF		x	x	x	United States
FlexShares STOXX Glbl ESG Impact ETF	x	x	x	x	Developed
FlexShares STOXX US ESG Impact ETF	x	x	x	x	United States
Global X CleanTech ETF			x	x	Developed
Global X Clean Water ETF					
Global X S&P 500® Catholic Values ETF	x	x	x	x	United States
Impact Shares NAACP Minority Empwrmt ETF	x	x	x	x	United States
Impact Shares Sus Dev Gls Glb Eq ETF	x	x	x	x	Developed
Impact Shares YWCA Women's Empwrmt ETF	x	x	x	x	United States
Inspire Faithward Large Cp Mmntm ESG ETF			x	x	United States
Inspire Faithward Mid Cap Mmntm ESG ETF			x	x	United States
Inspire International ESG ETF	x		x	x	Developed
Inspire Small/Mid Cap Impact ETF	x	x	x	x	United States
Invesco Global Clean Energy ETF	x	x	x	x	Developed
Invesco MSCI Sustainable Future ETF	x	x	x	x	Developed
Invesco Real Assets ESG ETF			x	x	Developed
Invesco US Large Cap Core ESG ETF			x	x	United States
Invesco MSCI Green Building ETF					
Invesco WilderHill Clean Energy ETF	x	x	x	x	United States
IQ Candriam ESG International Equity ETF			x	x	Developed
IQ Candriam ESG US Equity ETF			x	x	United States
iShares ESG Advanced MSCI EAFE ETF			x	x	Developed
iShares ESG Advanced MSCI EM ETF			x	x	Em. Markets

iShares ESG Aware MSCI EAFE ETF	x	x	x	x	Developed
iShares ESG Aware MSCI EM ETF	x	x	x	x	Em. Markets
iShares ESG Aware MSCI USA ETF	x	x	x	x	United States
iShares ESG Aware MSCI USA Small-Cap ETF	x	x	x	x	United States
iShares ESG MSCI USA Leaders ETF	x	x	x	x	United States
iShares Global Clean Energy ETF	x	x	x	x	Developed
iShares MSCI Global Impact ETF	x	x	x	x	Developed
iShares MSCI KLD 400 Social ETF	x	x	x	x	United States
iShares MSCI USA ESG Select ETF	x	x	x	x	United States
iShares® ESG Advanced MSCI USA ETF			x	x	United States
iShares® ESG MSCI EM Leaders ETF			x	x	Em. Markets
iShares® ESG Screened S&P 500 ETF			x	x	United States
iShares® ESG Screened S&P Mid-Cap ETF			x	x	United States
iShares® ESG Screened S&P Small-Cap ETF			x	x	United States
KraneShares MSCI China Clean Tech ETF	x	x	x	x	Asia Pacific
KraneShares MSCI China ESG Leaders ETF			x	x	Asia Pacific
Nuveen ESG Emerging Markets Equity ETF	x	x	x	x	Em. Markets
Nuveen ESG Intl Dev Mkts Eq ETF	x	x	x	x	Developed
Nuveen ESG Large-Cap ETF	x	x	x	x	United States
AVDR US LargeCap ESG ETF			x		
Nuveen ESG Large-Cap Growth ETF	x	x	x	x	United States
Nuveen ESG Large-Cap Value ETF	x	x	x	x	United States
Nuveen ESG Mid-Cap Growth ETF	x	x	x	x	United States
Nuveen ESG Mid-Cap Value ETF	x	x	x	x	United States
Nuveen ESG Small-Cap ETF	x	x	x	x	United States
PIMCO RAFI ESG US ETF			x	x	United States
SmartETFs Sustainable Energy II ETF			x	x	United States
SPDR® Kensho Clean Power ETF	x	x	x	x	United States
SPDR® S&P 500 Fossil Fuel Rsrv Free ETF	x	x	x	x	United States
SPDR® S&P 500® ESG ETF			x	x	United States
TrueShares ESG Active Opportunities ETF			x	x	United States
US Vegan Climate ETF			x	x	United States
VanEck Vectors Environmental Svcs ETF	x	x	x	x	United States
Stance Equity ESG Large Cap Core ETF					
Vanguard ESG International Stock ETF	x	x	x	x	Developed
Vanguard ESG US Stock ETF	x	x	x	x	United States
WisdomTree Emerging Markets ESG ETF	x	x	x	x	Em. Markets
SPDR® MSCI EAFE Fossil Fuel Free ETF		x	x		
SPDR® MSCI Em Mkts Fossil Fuel Free ETF		x	x		
WisdomTree International ESG	x	x	x	x	Developed
WisdomTree U.S. ESG ETF	x	x	x	x	United States
VanEck Vectors Social Sentiment ETF					
Xtrackers MSCI ACWI ex USA ESG LdrsEqETF	x	x	x	x	Developed
Xtrackers MSCI EAFE ESG Leaders Eq ETF	x	x	x	x	Developed
Xtrackers MSCI EMs ESG Leaders Eq ETF	x	x	x	x	Em. Markets
Xtrackers MSCI USA ESG Leaders Eq ETF	x	x	x	x	United States
Xtrackers S&P 500 ESG ETF	x	x	x	x	United States

Appendix B: Benchmark matching with ESG ETFs

This list contains all of the 86 ESG ETFs and the benchmark / parent index with which they were matched. Benchmark / parent indeces were selected based on ESG ETF prospectus information or information gathered from the ETF manager's website. Given the nature of some of the ETF's investments not all of the ESG ETFs were assigned a benchmark / parent index. These ETFs were also not included in the benchmark return calculation in Table 2.

Name	Benchmark / parent
Adasina Social Justice All Cp Glbl ETF	n.a.
ALPS Clean Energy ETF	n.a.
American Century Sustainable Equity ETF	S&P 500
American Conservative Values ETF	n.a.
AVDR US LargeCap ESG ETF	US Large cap index
Change Finance US LgCp FossilFuel Fr	US Large cap index
Changebridge Capital Sustainable Eq ETF	S&P 500
ClearBridge Dividend Strategy ESG ETF	S&P 500
ClearBridge Large Cap Growth ESG ETF	Russell 1000 growth index
Columbia Sustainable Intl Eq Inc ETF	MSCI World ex USA
Columbia Sustainable US Equity Inc ETF	MSCI USA index
Direxion MSCI USA ESG Ldrs vs Lggds	MSCI USA index
Ecofin Global Water ESG	n.a.
Etho Climate Leadership US ETF	n.a.
First Trust EIP Carbon Impact ETF	n.a.
First Trust NASDAQ® ABA Community Bk ETF	NASDAQ
FlexShares STOXX Glbl ESG Impact ETF	STOXX Global 1800
FlexShares STOXX US ESG Impact ETF	STOXX USA 900
Global X Clean Water ETF	n.a.
Global X CleanTech ETF	n.a.
Global X S&P 500® Catholic Values ETF	S&P 500
Impact Shares NAACP Minority Empwrmt	Morningstar US Large-Mid Cap index
Impact Shares Sus Dev Gls Glb Eq ETF	Morningstar Global markets Large-Mid Index
Impact Shares YWCA Women's Empwrmt	Morningstar US Large-Mid Cap index
Inspire Faithward Large Cp Mmntm ESG	n.a.
Inspire Faithward Mid Cap Mmntm ESG	USA Mid cap index
Inspire International ESG ETF	n.a.
Inspire Small/Mid Cap Impact ETF	US Small cap / US Mid cap
Invesco Global Clean Energy ETF	n.a.
Invesco MSCI Green Building ETF	MSCI ACWI IMI
Invesco MSCI Sustainable Future ETF	MSCI ACWI IMI
Invesco Real Assets ESG ETF	n.a.
Invesco US Large Cap Core ESG ETF	USA Large cap index
Invesco WilderHill Clean Energy ETF	n.a.
IQ Candriam ESG International Equity	Solactive GBS Developed markets ex North America Large & Mid Cap Index
IQ Candriam ESG US Equity ETF	Solactive GBS United States Large & Mid Cap Index
iShares ESG Advanced MSCI EAFE ETF	Solactive GBS United States Large & Mid Cap Index
iShares ESG Advanced MSCI EM ETF	MSCI Emerging markets index

iShares ESG Aware MSCI EAFE ETF	MSCI EAFE Index
iShares ESG Aware MSCI EM ETF	MSCI Emerging markets index
iShares ESG Aware MSCI USA ETF	MSCI USA index
iShares ESG Aware MSCI USA Small-Cap	MSCI USA Small cap index
iShares ESG MSCI USA Leaders ETF	MSCI USA index
iShares Global Clean Energy ETF	S&P Global BMI
iShares MSCI Global Impact ETF	MSCI ACWI Index
iShares MSCI KLD 400 Social ETF	MSCI USA IMI
iShares MSCI USA ESG Select ETF	MSCI USA index
iShares® ESG Advanced MSCI USA ETF	MSCI USA index
iShares® ESG MSCI EM Leaders ETF	MSCI Emerging markets index
iShares® ESG Screened S&P 500 ETF	S&P 500
iShares® ESG Screened S&P Mid-Cap	S&P Midcap 400
iShares® ESG Screened S&P Small-Cap	S&P Smallcap 600
KraneShares MSCI China Clean Tech ETF	MSCI China Index
KraneShares MSCI China ESG Leaders	MSCI China Index
Nuveen ESG Emerging Markets Equity	MSCI Emerging markets index
Nuveen ESG Intl Dev Mkts Eq ETF	MSCI EAFE Index
Nuveen ESG Large-Cap ETF	MSCI USA index
Nuveen ESG Large-Cap Growth ETF	MSCI USA growth Index
Nuveen ESG Large-Cap Value ETF	MSCI USA index
Nuveen ESG Mid-Cap Growth ETF	MSCI USA Mid-cap growth index
Nuveen ESG Mid-Cap Value ETF	MSCI USA Mid-cap value index
Nuveen ESG Small-Cap ETF	MSCI USA Small-cap index
PIMCO RAFI ESG US ETF	FTSE RAFI US 1000 Index
SmartETFs Sustainable Energy II ETF	MSCI World
SPDR® Kensho Clean Power ETF	n.a.
SPDR® MSCI EAFE Fossil Fuel Free	MSCI EAFE Index
SPDR® MSCI Em Mkts Fossil Fuel Free	MSCI Emerging markets index
SPDR® S&P 500 Fossil Fuel Rsrv Free	S&P 500
SPDR® S&P 500® ESG ETF	S&P 500
Stance Equity ESG Large Cap Core ETF	S&P 500 / Russell 1000
TrueShares ESG Active Opportunities ETF	USA Large cap index
US Vegan Climate ETF	n.a.
VanEck Vectors Environmental Svcs ETF	n.a.
VanEck Vectors Social Sentiment ETF	n.a.
Vanguard ESG International Stock ETF	FTSE Global All Cap ex US Index
Vanguard ESG US Stock ETF	FTSE USA All Cap Index
WisdomTree Emerging Markets ESG ETF	n.a.
WisdomTree International ESG	n.a.
WisdomTree U.S. ESG ETF	n.a.
Xtrackers MSCI ACWI ex USA ESG LdrsEqETF	MSCI ACWI ex USA Index
Xtrackers MSCI EAFE ESG Leaders Eq	MSCI EAFE Index
Xtrackers MSCI EMs ESG Leaders Eq	MSCI Emerging markets index
Xtrackers MSCI USA ESG Leaders Eq	MSCI USA index
Xtrackers S&P 500 ESG ETF	S&P 500
Xtrackers S&P MidCap 400 ESG ETF	S&P Midcap 400
Xtrackers S&P SmallCap 600 ESG ETF	S&P Smallcap 600

Appendix C: Overview of Benchmark ETFs

The following list contains an overview of the 45 ETFs that were selected to form the benchmark universe.

Benchmark	No.	ETF
FTSE Global All Cap ex US Index	1	Vanguard Total International Stock Index Fund ETF
FTSE RAFI US 1000 Index	1	Invesco FTSE RAFI US 1000 ETF
Morningstar US Large-Mid Cap index	1	iShares Morningstar U.S. Equity ETF
MSCI ACWI ex USA Index	1	iShares MSCI ACWI ex U.S. ETF
MSCI ACWI ex USA Index	2	SPDR MSCI ACWI ex-US ETF
MSCI ACWI IMI	1	SPDR MSCI ACWI IMI ETF US
MSCI ACWI Index	1	iShares MSCI ACWI ETF
MSCI ACWI Index	2	Xtrackers MSCI AC World Index UCITS ETF DR
MSCI ACWI Index	3	Hanwha ARIRANG SYNTH-MSCI AC World ETF H-Equity Derivatives
MSCI ACWI Index	4	SPDR MSCI ACWI UCITS ETF
MSCI China Index	1	iShares MSCI China ETF
MSCI China Index	2	Amundi ETF MSCI China UCITS ETF
MSCI China Index	3	X-trackers MSCI China Index UCITS ETF
MSCI EAFE Index	1	ISHARES MSCI EAFE ETF
MSCI EAFE Index	2	BMO MSCI EAFE Index ETF
MSCI EAFE Index	3	Hanwha ARIRANG Synth-MSCI EAFE ETF H
MSCI Emerging markets index	1	iShares MSCI Emerging Markets ETF
MSCI Emerging markets index	2	MSCI EM Power Buffer ETF
MSCI Emerging markets index	3	Amundi ETF MSCI Emerging Markets UCITS ETF - A
MSCI Emerging markets index	4	Xtrackers MSCI Emerging Markets Index UCITS ETF
MSCI USA index	1	Amundi ETF MSCI USA UCITS ETF
MSCI USA index	2	Xtrackers MSCI USA Index UCITS ETF
MSCI USA index	3	iShares MSCI USA UCITS ETF
MSCI USA Small cap index	1	iShares MSCI USA Small Cap UCITS ETF
MSCI USA Small cap index	2	Invesco PureBeta MSCI USA Small Cap Portfolio
MSCI USA Small-cap index	3	iShares MSCI USA Small Cap UCITS ETF
MSCI World	1	Amundi ETF MSCI World UCITS ETF
MSCI World	2	Xtrackers MSCI World Index UCITS ETF
MSCI World	3	HSBC MSCI World UCITS ETF

MSCI World	4	iShares MSCI World ETF
MSCI World ex USA	1	iShares Core MSCI International Developed Markets ETF
Russell 1000 growth index	1	iShares Russell 1000 Growth ETF
Russell 1000 growth index	2	Vanguard Russell 1000 Growth ETF
S&P 500	1	SPDR S&P 500 ETF Trust
S&P 500	2	iShares Core S&P 500 ETF
S&P 500	3	SPDR Portfolio S&P 500 ETF
S&P 500	4	Vanguard S&P 500 ETF
S&P Midcap 400	1	iShares Core S&P Mid-Cap ETF
S&P Midcap 400	2	SPDR S&P Midcap 400 ETF Trust
S&P Midcap 400	3	Vanguard S&P Mid-Cap 400 ETF
S&P Smallcap 600	1	iShares Core S&P Small-Cap ETF
S&P Smallcap 600	2	SPDR Portfolio S&P 600 Small Cap ETF
S&P Smallcap 600	3	Vanguard S&P Small-Cap 600 ETF
Solactive GBS Developed markets ex North America Large & Mid Cap Index	1	Goldman Sachs MarketBeta International Equity ETF
Solactive GBS United States Large & Mid Cap Index	1	Goldman Sachs MarketBeta US Equity ETF

Appendix D: Data availability of benchmark ETFs

The list consists of all 45 Benchmark ETFs that were selected through Morningstar Direct. "TR" is an indication of Thomson Reuters ESG score availability; "MS1" and "MS2" indicate availability of Morningstar Sustainability ratings. "Returns" indicate return availability, and "geography" shows which factors were used to calculate the fund's abnormal returns. Information on expense ratios was available for each ETF.

Name	TR	MS1	MS2	Returns	Geography
Amundi ETF MSCI China UCITS ETF					Asia Pacific
Amundi ETF MSCI Emerging Markets UCITS ETF					Em. Markets
Amundi ETF MSCI USA UCITS ETF					United States
Amundi ETF MSCI World UCITS ETF					Developed
BMO MSCI EAFE Index ETF		X	X		Developed
Goldman Sachs MarketBeta International Equity ETF			X		Developed
Goldman Sachs MarketBeta US Equity ETF			X		United States
Hanwha ARIRANG SYNTH-MSCI AC World ETF H-Equity Derivatives			X		Developed
Hanwha ARIRANG Synth-MSCI EAFE ETF H			X		Developed
HSBC MSCI World UCITS ETF		X	X		Developed
Invesco FTSE RAFI US 1000 ETF	X	X	X	X	United States
Invesco PureBeta MSCI USA Small Cap Portfolio		X	X		United States
iShares Core MSCI International Developed Markets ETF	X	X	X	X	Developed
iShares Core S&P 500 ETF	X	X	X	X	United States
iShares Core S&P Mid-Cap ETF	X	X	X	X	United States
iShares Core S&P Small-Cap ETF	X	X	X	X	United States
iShares Morningstar U.S. Equity ETF	X	X	X	X	United States
iShares MSCI ACWI ETF	X	X	X	X	Developed
iShares MSCI ACWI ex U.S. ETF	X	X	X	X	Developed
iShares MSCI China ETF	X	X	X	X	Asia Pacific
ISHARES MSCI EAFE ETF	X	X	X	X	Developed
iShares MSCI Emerging Markets ETF	X	X	X	X	Em. Markets
iShares MSCI USA Small Cap UCITS ETF		X	X		United States
iShares MSCI USA Small Cap UCITS ETF		X	X		United States
iShares MSCI USA UCITS ETF		X	X		United States
iShares MSCI World ETF	X	X	X	X	Developed
iShares Russell 1000 Growth ETF	X	X	X	X	United States
MSCI EM Power Buffer ETF					Em. Markets
SPDR MSCI ACWI ex-US ETF	X	X	X	X	Developed
SPDR MSCI ACWI IMI ETF US	X	X	X	X	Developed
SPDR MSCI ACWI UCITS ETF		X	X		Developed
SPDR Portfolio S&P 500 ETF	X	X	X	X	United States
SPDR Portfolio S&P 600 Small Cap ETF	X	X	X	X	United States
SPDR S&P 500 ETF Trust	X	X	X	X	United States
SPDR S&P Midcap 400 ETF Trust	X	X	X	X	United States
Vanguard Russell 1000 Growth ETF	X	X	X	X	United States
Vanguard S&P 500 ETF		X	X		United States

Vanguard S&P Mid-Cap 400 ETF	X	X	X	X	United States
Vanguard S&P Small-Cap 600 ETF	X	X	X	X	United States
Vanguard Total International Stock Index Fund ETF	X	X	X	X	Developed
Xtrackers MSCI AC World Index UCITS ETF DR		X	X		Developed
X-trackers MSCI China Index UCITS ETF		X	X		Asia Pacific
Xtrackers MSCI Emerging Markets Index UCITS			X		Em. Markets
Xtrackers MSCI USA Index UCITS ETF			X		United States
Xtrackers MSCI World Index UCITS ETF			X		Developed

Appendix E: Fama and French robustness check

Financial performance – Fama and French abnormal returns

	Fama and French abnormal returns				
ER	0.72 (0.85)	0.59 (1.21)	-0.17 (-0.78)	-0.027 (-0.09)	0.63 (0.93)
AUM		-0.017 (-0.59)	-0.00067 (-0.17)	0.0052 (1.27)	-0.0057 (-0.40)
TR			0.0000075 (0.23)		
MS1				0.000057 (0.54)	
MS2					0.0027*** (2.99)
Benchmark		0.0035 (1.34)	0.0017** (2.07)	0.00079 (1.28)	-0.0054* (-1.68)
Fixed effects	Time	Time	Time	Time	Time
Sample	ESG	ESG + benchmarks	ESG + benchmarks	ESG + benchmarks	ESG + benchmarks
N	71	92	71	71	92
Adj. R^2	0.16	0.19	0.11	0.098	0.093

Note: Table output shows coefficients and z- statistics within brackets. Standard Errors were clustered at the ETF level; ER is the expense ratio as acquired through CRSP, AUM represents the ETF's Assets Under Management, TR, MS1 and MS2 represent ETF sustainability scores. Benchmark is a dummy variable which is 1 when it concerns one of the benchmark ETFs and 0 when it concerns one of the ESG ETFs.

** p < 0.1 ** p <0 .05 *** p < 0.01*

Appendix F: ESG ETF Flow performance relationship per year

	Lagged returns		
Year	One period	Two periods	Three periods
2010	-0.67 (-0.62)	0.89 (1.16)	-0.69 (-0.79)
2011	-0.26 (-0.89)	0.089 (0.59)	0.59** (2.08)
2012	0.26 (0.53)	0.27** (2.14)	0.049 (0.10)
2013	1.023* (1.81)	-0.072 (-0.24)	-0.59 (-1.22)
2014	1.15*** (3.10)	0.12 (0.28)	0.19 (0.76)
2015	0.91 (1.21)	-0.036 (0.11)	-0.266 (-1.36)
2016	0.28 (0.83)	-0.078 (-0.23)	-0.30 (-1.01)
2017	0.95 (1.28)	0.36 (0.78)	-1.21 (-1.21)
2018	0.26 (1.27)	-0.32 (-0.77)	0.40 (1.06)
2019	0.90** (2.34)	1.16** (2.23)	0.94*** (2.80)
2020	0.91*** (3.12)	0.45 (0.82)	0.050 (0.13)

*Note: Table output shows coefficients and z- statistics within brackets. Standard errors were clustered at the ETF level; Flow is calculated as 100 * ($(AUM_t * return_t) – AUM_{t-1})/AUM_t$,, the return variables are the ETF's pre fee returns lagged by one, two or three time periods.*

** p < 0.1 ** p < 0.05 *** p < 0.01*