Sustainable Investing: Examining the Returns from ESG Scores and ESG Momentum

Erasmus University Rotterdam
MSc Financial Economics

Abstract

This paper studies the effects of Environmental, Social, and Governance (ESG) scores and the year-on-year change in ESG scores (ESG momentum) on yearly excess stock returns. The study focusses on the European equity market from 2002 to 2021, using Refinitiv's Asset4 ESG ratings. An overall significant negative effect is found of ESG scores on firm performance, both for the standalone scores as well as for momentum. This effect is mainly driven by the Environmental score. When reflecting on the first two pandemic years, it has been found that during 2019, high ESG rated stocks have not offered a safe haven for investors, whereas during 2020, this did happen to be the case. It seems investors are willing to pay a premium to make ethical choices and to contribute to a more sustainable and social world.

Eva Haro Alfaro 476417
Supervisor: Prof. dr. O.W. Steenbeek
Second assessor: Dr. Jan Lemmen
Date Final version: 10/02/23

ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Contents

1. Introduction 3
2. Literature review and hypothesis development 5
2.1 Corporate Social Responsibility 5
2.2 ESG investing 6
2.3 ESG rating firms 8
2.4 Do ESG investments under- or outperform? 10
2.5 Rationale 12
2.6 A Safe Haven? 13
2.7 Prospect theory versus ESG momentum 15
3. Data & Methodology 18
3.1 Refinitiv Database 18
3.2 Data Collection 19
3.3 Empirical Design 23
3.3.1 First Hypothesis 23
3.3.2 Second Hypothesis 24
3.3.3 Third Hypothesis 25
3.3.4 Fourth Hypothesis 26
4. Empirical results 28
4.1 The relationship between ESG ratings and excess returns 28
4.2 The performance of high ESG rated firms during the pandemic 32
4.3 Are upgrades rewarded more than downgrades are being punished? 34
4.4 Is ESG momentum a better predictor? 37
5. Discussion 39
6. Limitations and future research 42
Appendix 44
References 50

1. Introduction

With increasing levels of emissions, an incredible growth of population, and consumption of natural resources faster than can be replenished, the environmental and social issues are becoming more and more pressing (United Nations, 2019). Fortunately, this development has not gone unnoticed by the financial sector, and it has gained the attention of a growing number of investors.

Traditional financial is based on deep-rooted theories such as Portfolio Theory (Markowitz, 1952), and the Efficient Market Hypothesis (Fama, 1970), which state that investors are rational and make decisions that satisfy return maximalization (or, risk minimalization). Hence, the traditional investor is purely driven by a risk-return trade-off. Since then, it has become clear that as human beings, we are not as "rational" as we tend to believe ourselves to be, and these fundamental models often do not do a very good job at capturing reality (Schoenmaker & Schramade, 2019). Investors seem to care about more than just return maximization, which is reflected in the increasing popularity of sustainable investing. Bloomberg (2021) reports that by the end of 2021, the global ESG assets under management had reached $37.8 trillion, and is even expected to reach $53 trillion by 2025, representing more than a third of the $140.5 trillion in total assets under management. During 2021, Europe accounted for half of the global ESG assets under management.

An increasing group of investors is abandoning return maximization as the only goal, and has started to consider a more ethical impact within the Environmental, Social, and Governance (ESG) spheres. This often happens using negative screening by excluding so-called sin-stocks, which means excluding firms or whole industries involved in unethical practices. Some investors take it even further by only investing in highly sustainable firms with a high ESG rating.

Besides the increased attention among investors, this topic has also gained popularity among academics. The amount of papers investigating the effect of sustainability in firms on performance is rising, even though there is yet to find an unambiguous relation between the two. Some studies report a significant non-negative or even positive relationship between the two, such as Friede et al. (2015), Hvidkjær (2017), Bennani et al. (2018), Pedersen et al. (2020).

Furthermore, ESG information has been linked to lower capital constraints (Cheng et al., 2014), the protection of shareholder rights (Ferell et al., 2016), and enhancing the corporate reputation (Brown et al., 2009).

On the other hand, other studies have reported a non-significant or even a negative relationship between high ESG ratings and performance, such as Renneboog et al. (2008), Halbritter and Dorfleitner (2015), Edelman (2019), and more. Zhong and Chen (2019) added to this that investors are willing to pay a premium for investing in sustainable firms.

Also during times of crisis, mixed results have been found. Some have found that sustainable firms act as a safe haven during periods of crises (Cornett et al., 2016; Lööf et al., 2022), while others report no relationship (Albuquerque et al., 2020; Bae et al., 2021). This ambiguity in sustainable investing research has lead to the following research question:

What is the effect of ESG activities on a firm's financial performance and does it offer a safe haven during a crisis period?

Many studies have been done regarding the relationship between ESG screening and firm performance. When evaluating ESG screening, this study takes both standalone ESG scores as well as the yearly change in ESG scores into account, instead of just examining one of the two, which adds to the existing literature. Furthermore, since the pandemic has been so recent, there is not a lot of literature on the effects of sustainable investing during Covid-19. Lastly, most literature on this topic, for "normal" times as well as crises, focus on the US market, while this study looks at the entire European market.

The remainder of this paper is organised as follows. First, an overview of the development of ESG ratings over time, as well as its general implications will be provided. Then, an overview of the current state of the literature on sustainable investing will be presented for "normal" times, as well as crises periods. Thereafter, Prospect Theory will be introduced and linked to the research on ESG momentum. In the next chapter, the data and methodology used in this study will be explained. The last three chapters will present the results, discuss the findings, and name some limitations and implications for future research.

2. Literature Review and Hypothesis Development

2.1 Corporate Social Responsibility

The concept of sustainable investing has evolved widely over the past years, and has known many different forms, names, and parties. Broadly, the evolution of sustainable investing or its corporate counterpart, Corporate Social Responsibility (CSR), can be divided in four different stages: ethical investing, early socially responsible investing, current socially responsible investing, and Environmental, Social and Governance investing. An overview of these developments and the corresponding timeframe is presented in Figure 1 (Fulton et al., 2012).

The first form of sustainable investing, ethical investing, was the earliest form of socially responsible investing (SRI). This way of investing was characterized by negative screening, or consciously choosing not to invest in particular companies or industries that did not align with personal norms and values, also referred to as value-driven screening. Rooted in several different religions, it became highly popular to seek alignment between investment and faith. Around 1960, SRI regained popularity and appeared as a new concept. Also being value-driven, this investment strategy incorporated ethical, social, and environmental aspects. These developments gained more momentum and popularity because of the UN 1992 Conference on Environment and Development (UNCED), which also resulted in stronger political support. From the late 1990s up until the present, SRI started moving towards its modern form. A shift from a strong value-driven approach to a more return-seeking strategy appeared. Environmental, social, and corporate governance factors started to be incorporated into portfolio formation, which led to a strategy which is actively pursuing returns. The current SRI uses both negative (value-driven) and positive (risk/return driven) screening techniques to combine both return maximization and morally acceptable investment decisions. The 'Responsible Investor" arose, who considers both financial factors, as well as environmental, social and governance (ESG) factors (Fulton et al., 2012).

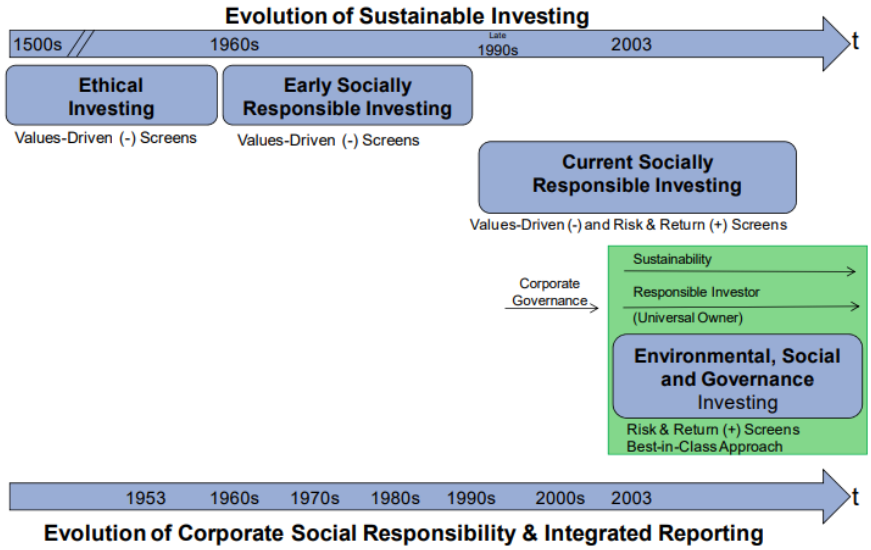

Figure 1. The Timeline of the Evolution of Sustainable Investing (Fulton et al., 2012).

2.2 ESG Investing

A more concrete definition of the until now unquantified concept of SRI emerged around 2003. The focused shifted even further towards a more risk and return driven focus regarding SRI. Multiple reports from the UNEP Finance Initiative showed that environmental, social and corporate governance issues affect long-term shareholder value, and in some cases, those effects may be profound (UNEP FI). Around 2006, the UN introduced the Principles for Responsible Investing (PRI), which, for the first time, talked about ESG investing as we know it. The Responsible Investor is now defined as the investor who incorporates ESG factors into his or her portfolio (Brochure PRI, 2021).

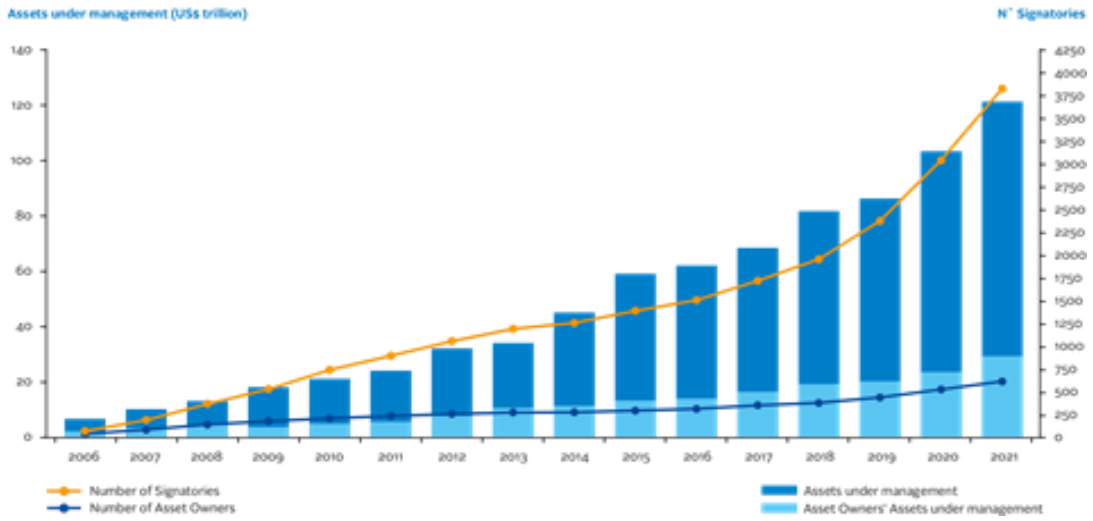

Figure 2. The PRI growth from 2006 until 2021 (Brochure PRI, 2021).

Since 2006, the PRI has seen consistent growth up until today (see Figure 2). As mentioned earlier, ESG stands for Environmental, Social, and Governance. Environmental captures aspects such as water consumption, waste, carbon emissions, etc. The social aspect deals with topics on diversity and human rights. Lastly, governance takes legal or regulatory aspects and anti-corruption into account (Amel-Zadeh & Serafeim, 2017). Over the years, sustainability aspects have become widely popular amongst investors and asset management firms. Institutional investors such as hedge funds are increasingly integrating sustainability into their portfolios. This development can be (partially) explained by the increase in the availability of information regarding SRI, the speed at which information can be shared, the highly increased awareness of environmental issues and the increased importance of transparency about corporate social responsibility (Lioui, 2018). The raised awareness on social inequalities, such as the underrepresentation of women, Black Lives Matter, and the MeToo movement, is increasingly visible in the world of finance and investments.

The effect of the incorporation of ESG factors into a portfolio on equilibrium market prices is inconclusive. Some investors simply maximize their mean-variance utility, and others use ESG scores as the main motivation for investments, excluding the so-called sin stocks. These are stocks that are part of industries such as tobacco, gambling, alcohol, and the weapon industry. This exclusionary practice can lead to major supply and demand imbalances (Zerbib, 2019).

However, the growing interest in sustainable investing is clear. In 2018 the total global sustainable investments amounted to $30.7 trillion, it has grown to about $35 trillion in 2020, and it is expected to grow to $50 trillion by 2025 (Bloomberg, 2022). Furthermore, the increase in ESG investing raises controversy regarding the classical asset pricing models. Investors apparently also derive utility from positive societal externalities, which means that investors seem to be willing to pay for impact (barber et al., 2021).

2.3 ESG Rating Firms

Over the last years, many different companies have entered the ESG rating market. Rating agencies such as Asset 4, Bloomberg, MSCI, Sustainalytics. Robeco and many more have developed their own ESG rating system (Figure 3). However, ESG is a highly heterogeneous space, and ESG rating companies play an increasingly important role in making investment decisions. This development comes with a large drawback; the different rating systems vary largely, which often leads to a drastic divergence in ratings for the same company (Billio et al., 2021).

Rating Agency	Abbrv.	Owner	Indicators
KLD	KL	MSCI	78
Sustainalytics	SA	Morningstar	163
Vigeo Eiris	VI	Moody's	38
RobecoSAM	RS	S&P Global	80
Asset4	A4	Refinitiv	282
MSCI	MS	MSCI	68

Figure 3. Overview of different rating firms (Billio et al., 2021)

There are three main sources that result in a divergence over ratings, according to Berg et al. (2019). The discrepancies in ratings are mostly driven by a difference in measurement, since rating agencies often measure the same attributes using different indicators (53%). Another, somewhat smaller source of discrepancy is a difference in scope, meaning that different agencies rely on a different set of attributes (44%). Lastly the least material source, which is the weight, meaning different agencies have different views on the relative importance of the attributes, also contribute (3%). In the same paper, Berg et al. (2019) show that the average correlation between the ESG scores of five popular ESG rating companies can differ from 42% to 74% (see Figure 4). This increases the challenge for investors to price ESG ratings correctly in the market (Hanicova & Vojtko. 2020)

Furthermore, the quality of the data used by ESG rating companies can be deficient because of the dependence on self-reporting and a lack of coverage. In a survey of institutional investors, BNP Paribas (2017) found that 55% of the respondents indicated that the greatest impediment to enter the ESG investing space is the lack of robust data.

	Asset4	RobecoSAM	Sustainalytics	Vigeo-Eiris	MSCI	KLD
Asset4	100%	62%	67%	69%	38%	42%
RobecoSAM	62%	100%	67%	70%	38%	44%
Sustainalytics	67%	67%	100%	71%	46%	53%
Vigeo-Eiris	69%	70%	71%	100%	42%	49%
MSCI	38%	38%	46%	42%	100%	53%
KLD	42%	44%	53%	49%	53%	100%

Figure 4. Correlation between ESG scores (Berg et al., 2019)

2.4 Do ESG Investments Under- or Outperform?

For investors and asset managers it is very beneficial to know whether ESG ratings have any predictive power on firm performance and share price development, and over the years, many researches have been done on assessing the performance of a high ESG rated firm. The results are, however, quite ambiguous. A highlight will be discussed in this section.

Prior studies have shown that ESG disclosures are associated with different economically relevant effects. Cheng et al. (2014) shows that ESG information is linked to lower capital constraints, and Dhaliwal et al. (2011) find a positive relation between ESG disclosures and lower costs of capital. Lins et al. (2017) found that firms with a higher CSR score outperformed others on more growth, profitability, and sales per employee, which in turn benefits shareholders as well. This is also found by Ferell et al. (2016), who also find a positive relationship between CSR activities and the protection of shareholder rights. Furthermore, Edmans (2011) found that higher employee satisfaction leads to higher shareholder returns. Hvidkjær (2017) also finds evidence that stocks with a high ESG rating exhibit high future returns. This is most strongly present in 1991 until 2004, while the high rated ESG stocks do not differ from benchmarks during 2005 until 2012. They do find some evidence that the returns have been higher again since 2012. Bennani et al. (2018) document a somewhat similar pattern; high ESG rated stocks did not outperform benchmarks during 2010 until 2013, but thereafter, it did became profitable. Pedersen et al. (2020) also find that high ESG rated stock outperform, but it does depend on whether the value of the ESG score is not fully priced in the market. The positive relation weakens when investors are willing to accept lower returns for having more sustainable stocks. Friede et al. (2015) study around 2,200 researches that cover the relationship between ESG rating and financial performance. They find that about 90% of the researches state a non-negative relationship, and the majority showed even a positive relationship. Buallay (2018) also looks at whether ESG reporting is associated with performance, and he looks at the effect on ROA, Tobin's Q, and ROE. Overall, he finds a significant positive relation between ESG disclosure and firm performance.

However, when measured individually, the results change. The environmental aspect positively affects the ROA and Tobin's Q, whereas the social pillar is negatively associated to all three models. Governance disclosure negatively affects the ROE and ROA, while it positively affects Tobin's Q. ESG disclosure also brings other advantages, such as increased transparency by showing links between financial performance and ESG information (Adams, 2017), it enhances the corporate reputation and creates a significant competitive advantage (Brown et al., 2009; Simnett et al., 2009).

On the contrary, many studies have also shown that there is no, or a negative relationship between incorporating ESG factors and performance. Friedman (1970) stated in his paper that ESG activities lead to agency problems between shareholders and managers, since the ESG activities come at the costs of the shareholders. Renneboog et al. (2008) supported this view, and found that SRI funds tend to underperform the benchmarks. According to them, the high ESG rated companies are overpriced, meaning investors have to pay a premium for making sustainable investment decisions. Another research from Halbritter and Dorfleitner (2015) also shows that there is no significant difference in performance for firms with a high or low ESG score. Edelman (2019) did a survey amongst 600 institutional investors, and found that most respondents would invest in ESG stocks, when given a lower rate of return. Some recent papers confirmed this finding, and stated that investors are willing to pay a premium when investing in sustainable firms (Zhong & Chen, 2019). There also exists quite some evidence on the outperformance of sin stocks relative to their benchmarks. Derwall et al. (2011) call this the shunned-stock hypothesis, which states that socially controversial stocks usually trade at a discount, since they are shunned by socially responsible investors. Multiple other papers show that the restriction of the investment universe for exclusionary practices could lead to diminished diversification, which comes at a cost for the responsible investor, lowering the expected portfolio returns, and might increase idiosyncratic risk (Hong & Kacperczyk, 2009; Hanicova & Vojtko, 2020). Statman and Glushkov (2009) find that investing in high ESG rated stocks is advantageous, however, this advantage is offset by the corresponding disadvantage of not investing in sin stocks.

2.5 Rationale

In conclusion, the existing literature is highly ambiguous, partly because of the different rating methodologies and financial metrics used to assess the financial performance (Berg et al., 2019). For both the under- and outperformance of high ESG rated stocks exist plausible reasons.

The theory that high ESG rated firms show a better performance finds its ground in the argument that the stock market generally underreacts to information related to ESG. This means that the effects of positive ESG-related events are not (fully) recognized in the market, which leads to certain firms being undervalued, and investing in these companies can yield abnormal returns. This so-called Underreaction Hypothesis is not new in financial literature; there exists a lot of evidence that stock markets tend to underreact in particular situations. For example, Bernard and Thomas (1989) prove this hypothesis regarding post-earnings announcement drift, and Jagadeesh and Titman (1993) describe the well-known momentum effect. More specifically, it could be reasonable to assume that intangibles are more difficult to value than tangibles, meaning that the stock market would undervalue certain intangibles. Prior literature points out evidence of underreaction for intangibles such as software development, advertising, and R&D costs (Edmans, 2011). In the same line of reasoning, ESG investments are often classified as intangibles, which makes it probable for the stock market to underreact. Another plausible reason for outperformance might be that sustainable investing has become increasingly popular recently, meaning an increasing demand for sustainable stocks, which in turn drives up the price of these stocks (Hvidkjær, 2017).

However, the demand effect might also be exactly the reason for underperformance of high ESG stocks. When a big group of investors focus on one group of stocks (high ESG stocks), and ignore another group of stocks (sin stocks), the sin stocks can become undervalued (Merton, 1987). While this leads to initial low returns, these stocks will thereafter show high returns relative to high ESG rated stocks, leading to the underperformance of highly sustainable firms. Also, sin stocks are often stocks from companies in industries such as weapons or tobacco, and these firms often practice conservative accounting, since these industries fall under scrutiny from regulators. Many

investors might not account for this, leading to underreaction and high returns of sin stocks (Hong & Kaperczyk, 2009). This leads to the first hypothesis that will be tested:

A higher ESG rating leads to higher stock returns.

2.6 A Safe Haven?

Prior literature has shown that known relations between factors can change drastically during times of crisis. As we are currently recovering from a major crisis period, the question arises what this time of the pandemic has done to the returns of sustainable investing.

During turmoil times, risk-averse investors often actively seek low risk and high performance safe-haven investments (Rubbaniy et al., 2021). Stocks with a higher ESG ranking are often considered to be safer during stock market turmoil, and also expected to exhibit greater potential for future recovery from a crisis (Lööf et al., 2022). Research on the financial crisis from 2008 shows that firms with a high social capital, measured by CSR intensity, were significantly less affected by the crisis than firms with low social capital, and had a better financial performance (Lins et al., 2017). Cornett et al. (2016) show that during the global financial crisis, the financial performance of U.S. banks' was positively related to their ESG score, which is consistent with the theory of a 'flight to quality'.

The COVID-19 pandemic has already been used widely by financial researchers to establish what kind of effect a crisis period like this has on the demand for sustainable investments, and the performance of high ESG stocks. For instance, Broadstock et al. (2021) look at the ESG performance in China before and during the pandemic, and show that high ESG rated portfolios generally outperform low ESG rated portfolios. They also find that a high ESG performance mitigates the financial risk during crisis. Albuquerque et al. (2020) state that high ESG rated stocks are more resilient during a crisis period and show for the first quarter of 2020 significantly higher returns, lower volatility, and higher trading volumes than other stocks in the same period.

However, using the same dataset as Albuquerque et al. (2020), Bae et al. (2021) report no evidence of a relation between CSR and stock returns during the crash period. Furthermore, equity funds with low ESG risk scores had positive investment inflows during the panic phase of COVID-19 and afterwards, while high risk ESG funds experienced sell-offs (Ferriani and Natoli, 2020). Studying the period of May 2017 until May 2020, Singh (2020) found that risk averse investors sought shelter in high ESG rated portfolios during the stock market collapse. The demand for sustainability during the crisis period significantly weakened the economic stress induced by the pandemic (Döttling & Kim, 2020). For investors, institutions, funds, and analysts, it is very relevant to know whether it could add significant value to shift wealth towards high ESG rated companies in times of crisis. This leads to the following hypothesis:

High ESG rated stocks have outperformed low ESG rated stocks during COVID-19.

2.7 Prospect Theory versus ESG Momentum

One of the most fascinating and upcoming areas in financial research is the quest for an explanation on how and why individuals make decisions. In 1953, Von Neumann and Morgenstem (1953) introduced the now most accepted model of rational choice; the expected utility theory. During the following years, the completeness of their theory has been increasingly questioned, and around 1980, another theory, rooted in behavioural economics, found its way through; prospect theory. This framework was formulated by Kahneman and Tversky (1979), and provided a more accurate description of how decisions are being made under risky situations (Edwards, 1996). According to this theory, agents are subject to several biases and violations with when it comes to expected utility theory. Many of these violations are explicitly predicted by prospect theory, and therefore, it forms a great improvement over expected utility theory. The substitute for the utility function, is the value function, and looks as follows (also see Figure 5):

$V(x) = x^{\alpha}$ if $x > 0$

$V(x) = -\lambda(-x^{\alpha})$ if $x < 0$

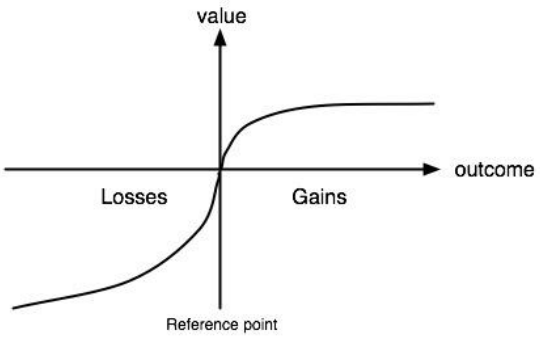

Figure 5. Prospect Theory Value Function (Kahneman and Tversky, (1979).

Prospect theory has three major characteristics:

i) Reference level dependence: an agent evaluates consequences in terms of deviations from a *reference level*, which usually is that person's *status quo* (Weber & Johnson, 2009).

ii) Gains and losses satiate: the value function is marginally decreasing, meaning that the value of a (both positive and negative) outcome has diminishing return characteristics. In the function mentioned above, this means $\alpha < 0$. Empirical studies estimate α to be equal to 0.88 (Weber & Johnson, 2009).

iii) Loss aversion: the value function is steeper than losses than for gains, meaning that losing \$10 results in more pain, than gaining \$10 results in pleasure. In the value function, λ captures this effect, and is typically estimated at 2.25. The λ being greater than one, means that losses weigh more heavily than gains (Weber & Johnson, 2009).

In conclusion, prospect theory shows that investors dislike financial losses more than they appreciate equivalent financial gains. When applying loss aversion to sustainable investing, you would expect investors to respond more negatively to a decline in a firm's sustainability score, than to respond positively to the same increase. However, Giese and Nagy (2018) performed an extensive research on this, and found evidence for the opposite. They report that MSCI ESG upgrades seem to outweigh downgrades, which stands in direct contrast to prospect theory. They also find that the ESG valuation curve was the steepest in the middle range of MSCI ESG scores, and relatively flat for very high and high ESG scores (see Figure 6) (Giese & Nagy, 2018). This leads to the third hypothesis:

An ESG rating upgrade is rewarded more than an ESG downgrade is punished.

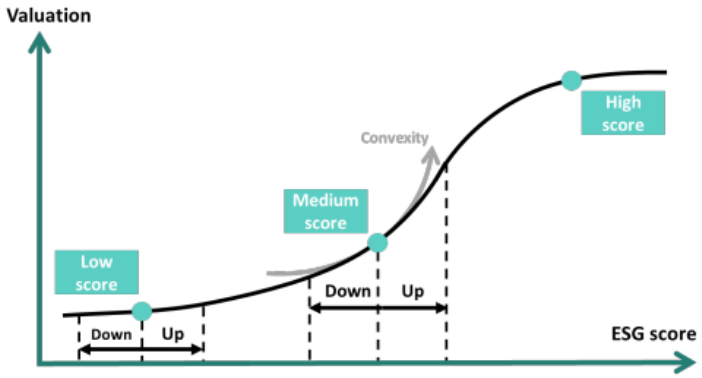

Figure 6. Stylized ESG-valuation Curve (Giese & Nagy, 2018).

The change in ESG rating over a year, is called ESG momentum. The financial value of ESG momentum is supported in prior literature, by for example Khan et al. (2016), Nagy et al. (2016), and Giese et al. (2019). While in earlier papers, standalone ESG score has not been a unambiguous predictor for financial performance, ESG momentum tends to do a better job. The authors of the papers mentioned above, showed a statistically significant predictive power of ESG momentum for stock returns. It may be a good financial indicator on its own right, and it could be useful to add this factor in portfolio construction (Giese et al., 2019). Giese and Nagy (2018) also show that MSCI ESG scores have shown correlations to quality factors, the low volatility factor, and the value factor, while ESG momentum scores have shown to be uncorrelated to any of the equity style factors. Therefore, ESG momentum can be quite straightforward to apply in portfolio construction, since there is little risk of suffering undesired factor biases when using ESG momentum. Furthermore, this methodology is rather short-term based, so the effect of new information can be captured, while ESG ratings itself focus on a more long-term value creation. This leads to the fourth and last hypothesis:

ESG momentum is a better predictor for performance, relative to standalone ESG scores.

These hypotheses lead to the following research question:

What is the effect of ESG activities on a firm's financial performance and does it offer a safe haven during a crisis period?

3. Data & Methodology

3.1 Refinitiv Database

The Asset4 database owned by Refinitiv has been rising in popularity, mainly because of recent criticism regarding the MSCI ESG rating methodology (Bouten et al., 2017). One concern they mention is that MSCI doesn't make a valid distinction between different industries when assigning weights to the different ESG factors. This is in contrast to Asset4, which customizes ESG metrics per firm, while accounting for industry differences. Another advantage of Asset4 is that the Refinitiv ESG scores range from 0 to 100, where a higher score renders a better sustainability performance. This method of rating lends itself to a rather precise investigation of changes in ESG scores. Therefore, this research will focus on the Refinitiv database for obtaining the ESG scores. Asset4 is one of the most comprehensive databases regarding ESG rating data. This database has been widely used in prior literature by, amongst others, Ferrell et al. (2016), Dyck et al. (2019), and Albuquerque et al. (2020). It contains data on over 10,000 firms across 76 countries, and covers around 80% of the global market capitalization (Drempetic et al., 2020). The scores are based on the self-reported information on the three different pillars of ESG, and calculated using publicly available company data as well as data provided by governmental institutions, NGOs, and trade unions. The three different pillars of ESG are further divided into subcategories, as can be seen in Figure 7. The ESG scores are based on information from sources such as annual reports, NGOs, and sustainability reports. The environmental pillar covers resource use, emissions, and innovation. The social pillar contains workforce, human rights, community, and product responsibility. And lastly, governance consists of management, shareholders, and CSR strategy. The database also contains a controversies score, which measures the exposure to environmental, social and governance controversies or negative events reflected in global media. When this score is included, these four pillars add up to an ESGC score (Refinitiv, 2022).

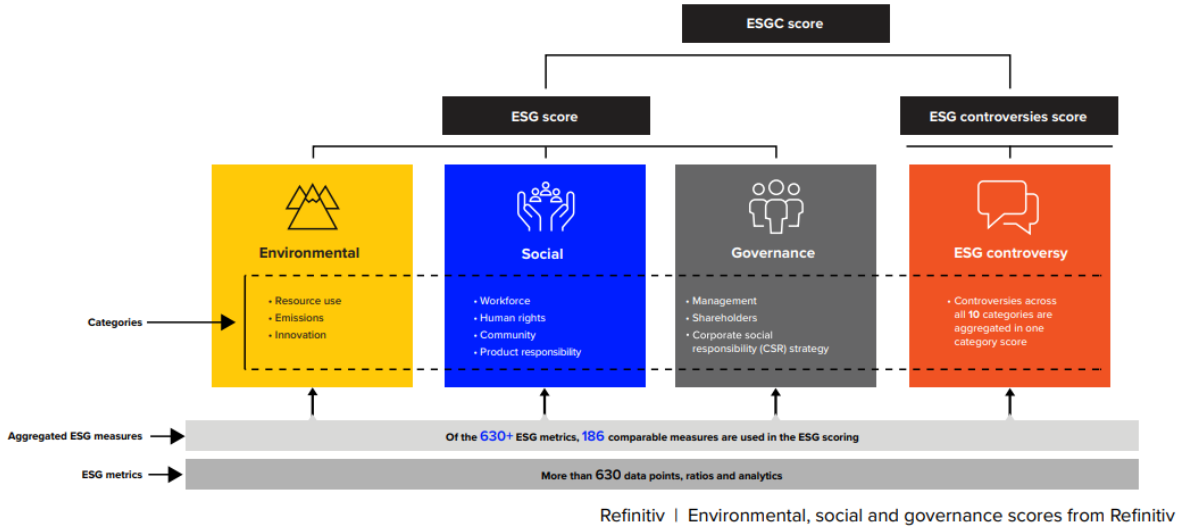

Figure 7. Asset 4 Pillars of ESG (Refinitiv, 2022).

3.2 Data Collection

This paper will focus on the European equity market from the beginning of 2002 until the end of 2021. The main variable of interest is the Refinitiv ESG score. These ratings are retrieved from the Thomson Reuter Refinitiv Asset4 database. Regarding the first hypothesis, apart from the overall ESG scores, the data on each contributing pillar is also retrieved from this database, and will be used to test for the individual effects. The ESG scores are updated annually. However, the ESG data disclosure is unstructured and can be published at any time of the year (Refinitiv ESG Scores Factsheet, 2022). For simplicity, it is therefore assumed that the scores are updated at the beginning of the year, and the effect is measured over that particular year.

Next, as is commonly done in financial researches, the data sample is narrowed down by excluding companies from the financial and the utilities industry (SIC codes 6,000-6,999 and 4,900-4,999 respectively), since these industries are subject to rather exceptional regulations, which decreases comparability to other firms.

The companies obtained from Asset4 are cross referenced with yearly stock prices. The yearly stock prices are converted into yearly logarithmic returns because stock returns presume normality, and logarithmic conversion will increase the returns' normality. To create excess returns, a yearly risk-free rate is subtracted from the yearly returns. A commonly used proxy for the European risk-free rate is the 3-month EURIBOR rate, which is also used in this research (ECB, 2022).

Regarding the second hypothesis on the COVID-19 crisis, some dummy variables are created to measure the effect during the pandemic. The Covid period in Europe is defined as the period starting January 1st 2020, and ending when the dataset for this research ends; December 31 2021. The Covid period will be further divided to measure the separate effects of both the Covid years in the sample.

From the ESG rating data, the variable ESG momentum is constructed. This is defined as the year-on-year change in ESG ratings, following Giese et al. (2019). However, the authors of this paper calculate the ESG momentum as the year-on-year percentage change, which leads to some methodological issues. The main limitation is outlined by the following example:
When the ESG score of firm A increases from 5 to 10, this leads to an increase of 100%. An increase in ESG score for firm B from 50 to 100, also leads to an increase of 100%, meaning both firms experience a similar ESG momentum, even though the increase of firm B is much more impressive. Also, using percentage change positively skews the results, since positive ESG momentum is not limited, while negative ESG momentum is capped at -100%. Lastly, using absolute values instead of percentages will ease the intuition and practicality behind the coefficients and results in general. The model will allow for a time lag of one year between the ESG rating change and the corresponding price reaction of the market. A one-year horizon is proven to be optimal for robust performance results for ESG momentum (Giese & Nagy, 2018).

Furthermore, accounting data are obtained from Datastream and are used to control for the usual firm characteristics, namely Tobin's q, assets, size, cash, leverage, dividend yield and historical volatility, following the methodology of Albuquerque et al. (2020). An overview of the definitions of all variables are presented in Table 1A in the appendix. After merging all the datasets, the sample consists of 2,584 distinct firms, which cover a total of 20 yearly observations, resulting

in a total of 51,680 firm-year return observations. Table 1 shows the summary statistics of the sample.

Table 1. Descriptive statistics

Variable	Observations	Mean	Std. Dev.	Minimum	Maximum
ESG Score variables					
Excess Return	38,300	0.015	0.399	-4.090	3.189
ESG Score	13,784	49.312	21.054	0.630	95.660
ESG Momentum	12,851	7.221	14.247	-16.220	61.790
E Pillar	12,633	48.342	25.461	0.030	99.220
S Pillar	13,773	51.333	24.739	0.340	98.630
G Pillar	13,784	50.461	22.573	0.570	99.330
ESG Momentum	12,851	7.221	14.247	-16.220	61.790
Control variables					
Tobin's q	36,759	1.216	0.976	0.104	7.077
Cash to Assets	37,400	0.104	0.113	0.000	0.702
Dividend Yield	26,849	3.278	2.227	0.250	15.200
Historical Volatility	37,332	0.353	0.168	0.096	1.166
Leverage	37,796	38.052	24.290	0.080	124.250
Size	34,310	4167.171	7799.796	36.600	48813.18

All control variables and the ESG momentum variable are winsorized at the 1% level in each tail, which takes care of extremely high values of skewness and/or kurtosis. The market capitalization however, is winsorized at the 5% level in each tail, since the kurtosis remained extremely high (around 600) after accounting for the 1% tails. After winsorizing, all variables have a skewness between -3 and 3, and a kurtosis around -10 to 10.

The yearly mean excess return over the last two decades is around 1.5%, with a standard deviation of 39.9%. This make sense, since the yearly stock price development for a firm can be quite large, which is reflected in the low minimum as well as the high maximum excess return. For a visual representation of the distribution, see Figure 8 in the appendix. The ESG Scores in the sample range from 0.63 until 95.66, with a mean of 49.31. The means of the separate pillars, E, S and G, are all centred around the mean of the ESG scores, respectively 48, 51 and 50. All four standard deviations are approximately equal as well, with the lowest of the ESG Score itself of 21, and the highest of the Environmental pillar of 25 (See Figure 9 to 12 in the appendix for the distributions of the variables). The ESG scores are thus not very stable over time; they rarely remain the same after a year. This can also be seen when looking at the ESG momentum statistics. The average yearly change in ESG Scores is 7.22, with a standard deviation of 14. The score upgrades are a lot higher than the score downgrades, with the maximum upgrade being 62, while the maximum downgrade is around 16, making this variable skewed to the right (see Figure 13 in the Appendix).

3.3 Empirical Design

3.3.1 First Hypothesis

To study the effect of ESG ratings on financial performance, multiple regressions will be run. The data is in panel form, which examines both group and time effects. It allows for controlling for variables that cannot be measured or observed, thus accounting for individual heterogeneity. These effects can be either fixed or random. A fixed effects model assumes that the entity´s error term and the predictor variables are correlated, and the fixed effects remove the effect of these time-invariant characteristics, so the net effect of the predictors can be assessed on the outcome variable. A random effects model assumes that the variation across entities is random and uncorrelated with the independent variables in the model (Torres-Reyna, 2007). To decide between a fixed or a random effects model, a Hausman test is done, which tests whether the unique errors are correlated with the regressors.

Hereafter, the first hypothesis that states: "A higher ESG rating leads to higher stock returns" will be tested using the following regressions:

$$R_{,i,t} - R_{F,\,t} = \alpha + ß_1\,\text{ESGscore}_{i,t} + ß_2\,\text{FirmControls}_{i,t} + \varepsilon_i \quad (1)$$

$$R_{,i,t} - R_{F,\,t} = \alpha + ß_1\,\text{E-score}_{i,t} + ß_2\,\text{S-score}_{i,t} + ß_3\,\text{G-score}_{i,t} + ß_4\,\text{FirmControls}_{i,t} + \varepsilon_i \quad (2)$$

Where $R_{,i,t} - R_{F,\,t}$ represents the stock return over the risk-free rate, $\text{ESGscore}_{i,t}$ shows the ESG rating for firm *i* on date *t,* $\text{E-score}_{i,t}$, $\text{S-score}_{i,t}$ and $\text{G-score}_{i,t}$ capture respectively the environmental, the social, and the governance score, and $\text{FirmControls}_{i,t}$ is a vector consisting of the control variables: Tobin's q, cash to assets, dividend yield, historical volatility, leverage, and size.

The coefficient of interest in Regression (1) is $ß_1$, which is expected to be positive.

This would mean that when the ESG score increases, the excess return increases as well, and this would support the first hypothesis. In Regression (2), it can be tested in which pillar the observed effect of $ß_1$ from Regression (1) is located most strongly.

3.3.2 Second Hypothesis

To test the second hypothesis, which states that high ESG-rated firms were associated with higher returns during the pandemic, the following regression will be used:

$$R_{i,t} - R_{F,t} = \alpha + ß_1 \text{HighESG}_t + ß_2 \text{CovidPeriod}_t + ß_3 \text{HighESG}_t * \text{CovidPeriod}_t + ß_4 \text{FirmControls}_{i,t} + \varepsilon_i \quad (3)$$

$$R_{i,t} - R_{F,t} = \alpha + ß_1 \text{HighESG}_t + ß_2 \text{CovidYear1}_t + ß_3 \text{HighESG}_t * \text{CovidYear1}_t + ß_6 \text{FirmControls}_{i,t} + \varepsilon_i \quad (4)$$

$$R_{i,t} - R_{F,t} = \alpha + ß_1 \text{HighESG}_t + ß_2 \text{CovidYear2}_t + ß_3 \text{HighESG}_t * \text{CovidYear2}_t + ß_6 \text{FirmControls}_{i,t} + \varepsilon_i \quad (5)$$

The variable $\text{HighESG}_{i,t}$ is a dummy variable equal to one when the ESG rating is in the top 50% of scores, and 0 when it is below. The CovidPeriod_t variable in Regression (3) is a dummy variable that takes on the value one from January 1 2020 to December 31 2021. The interaction term reflects the expectation is that the effect of high ESG scores on the excess return will change, given that we are in the Covid period. Because of the interaction effect, the interpretation of the beta will change. However, the expectation is that there will be a significant and positive effect of the high ESG rated companies in Covid period, on the excess returns. For Regressions (4) and (5), the dummy variables CovidYear1_t and CovidYear2_t are used. They take on the value one in respectively 2019 and 2020.

This way, it can be measured in which year the effect of the $CovidPeriod_t$ dummy is most strongly present. Interaction effects are also added, for the same reason as in Regression (3).

3.3.3 Third Hypothesis

To test the third hypothesis and assess whether ESG score upgrades are rewarded more than ESG downgrades are being punished, multiple regressions will be used. Firstly, a dummy variable is created with value one when ESG momentum is positive, hence, there is an upgrade, and zero otherwise. This way it can be tested if an ESG upgrade is indeed rewarded by investors, and if this is the case, whether it is rewarded more than downgrades are being punished. This leads to the following regression:

$$R_{i,\,t+1} - R_{f,\,t+1} = \alpha + ß_1 U_t + ß_2\, FirmControls_{i,t} + \varepsilon_i \qquad (6)$$

Where U_t represents the above mentioned dummy variable Upgrades at time *t* and $R_{i,\,t+1} - R_{f,\,t+1}$ shows the excess stock returns at time *t+1,* accounting for the one year lag for investors to process the new information and incorporate it into the share price. Note that the Downgrades variable is omitted in the regression, for reasons of multicollinearity. With this model, the reaction of an ESG score upgrade relative to a downgrade can be measured. The expectation is that $ß_1$ is positive, meaning that relative to the downgrade group, being in the upgrade group leads to a higher excess return.

To further investigate this effect, nine groups are created from the ESG momentum variable. The reference group contains the ESG score changes from -0.5 to 0.5. Then, with steps of 10 points, the other groups are formed. Since the maximum downgrade is about 16, there are merely two groups below the reference group.

The maximum upgrade is around 60, meaning that there are six groups above the reference group. See Table 2A in the appendix for the exact distribution along all nine groups. This leads to the following equation:

$$R_{i,\,t+1} - R_{f,\,t+1} = \alpha + ß_1 G1_t + ß_2 G2_t + ß_3 G3_t + ß_4 G4_t + ß_5 G5_t + ß_6 G6_t + ß_7 G7_t + ß_8 G8_t + ß_9 G9_t + ß_{10} \text{FirmControls}_{i,t} + \varepsilon_i \quad (7)$$

Where $G1_t$ captures the lowest ESG momentum values, and $G9_t$ the highest ones at time *t,* $R_{i,\,t+1} - R_{f,\,t+1}$ shows the excess stock returns at time *t+1,* again accounting for the assumed one year lag. Group 3 serves as the reference group, and is omitted in the regressions for reasons of multicollinearity. This means that the interpretation of the other variables will be relative to the third group. Since the values in this group are closely centred around zero, the other variables are compared to a score change of about zero, which eases interpretation. The expectation is that the positive effect on the excess return will increase when moving from large downgrades to large upgrades. This reflects the skewness regarding the investor valuation of ESG rating changes.

3.3.4 Fourth Hypothesis

To gain a better understanding on how investors treat changes in ESG scores, ESG momentum is introduced into the model. The following regression is used to measure the effect of ESG momentum on financial performance. This way, the fourth hypothesis can be tested, which states that ESG momentum is a better predictor for performance, relative to standalone ESG scores.

$$R_{i,\,t+1} - R_{f,\,t+1} = \alpha + ß_1 \text{ESGmomentum}_t + ß_2 \text{FirmControls}_{i,t} + \varepsilon_i \quad (8)$$

Where ESGmomentum a continuous variable representing the year-on-year change in the ESG score. Following the fourth hypothesis, the expectation is that $ß_1$ has a greater magnitude than the $ß_1$ from Model 1, since ESG momentum has shown to be a more stable and better predictor of returns compared to ESG scores.

4. Empirical results

4.1 The Relationship between ESG Ratings and Excess Returns

To test the first hypothesis, which states that a higher ESG score leads to a better performance, Regression (1) is used. Table 2 presents the results of regressing the excess returns on firm's ESG rating and other firm characteristics. In the first column, ESG rating is the only independent variable used. Next, both a random effects and a fixed effects model regression is done, and using the stored coefficients, a Hausman test is performed. The coefficients of the fixed and the random model differ significantly, with a p-value < 0.001. This means that a fixed effects model suits the data best. In the second column of Table 2, fixed effects are added. Lastly, in column 3, the control variables are added. Standard errors are clustered and robust to heteroskedasticity.

Table 2. Panel regression of yearly excess returns on Refinitiv ESG score

This table shows the effect of standalone ESG scores on firm performance including firm controls. The ESG score variable is highly significant and negatively related to firm performance.

	(1)	(2)	(3)
Dependent variable	Excess return	Excess return	Excess return
ESG	-0.00017	-0.00066**	-0.00119***
	(-0.98)	(-2.42)	(-3.79)
Tobin's q			0.08032***
			(7.46)
Cash to Assets			0.46929***
			(5.72)
Dividend Yield			-0.07893***
			(-20.60)
Historical Volatility			-0.96111***
			(_22.09)
Leverage			-0.00175***
			(-5.34)
Size			0.0000***
			(3.11)
Fixed Effects	No	Yes	Yes
R^2	0.000	0.000	0.2518

This table reports the results of regressions of the excess returns on firms' ESG score under several specifications: without firm controls (specification 1), with fixed effects (specification 2), and with fixed effects and firms controls (specification 3). Control variables are winsorized at the 1% level in each tail (except for Size at the 5% level). Standard errors are heteroscedasticity robust. The regression constant is not reported for brevity. The t-statistics are reported in parentheses. Table A1 in the appendix defines all the variables used. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

When not accounting for fixed effects, the effect of the ESG score on the excess returns appears insignificant. However, when adding the fixed effect, the coefficient changes, and is significant at the 5% level. Adding the control variables to the model increases the R^2 from < 0.001 to 0.2518, which makes the third specification the most appropriate one for interpreting the coefficient of the ESG score. After adding the control variables, it can be seen that the relationship between ESG rating and yearly excess returns turns out to be negative, in contrast to the first hypothesis. When a firm's ESG score goes up by one point, the firm's yearly excess returns lower by 0.12%, and this is significant at the 1% level.

Next, Regression 2 is used to test for the individual effects of the separate pillars of the ESG score. The Hausman test shows that as for ESG score, also for model (2), a fixed effects model should be used.

Table 3. Panel regression of yearly excess returns on the Environmental, Social, and Governance pillar

This table shows the effect of the three separate pillars of ESG scores on firm performance including firm controls. Only the Environmental score shows a significant negative effect on firm performance, whereas the Social and Governance score do not impact the returns.

	(1)	(2)	(3)
Dependent variable	Excess return	Excess return	Excess return
E	-0.00077***	-0.00154***	-0.00095***
	(-4.01)	(-4.66)	(-3.08)
S	0.00069***	0.00109***	-0.00011
	(2.80)	(3.26)	(-0.36)
G	-0.00005	0.00011	0.0000
	(-0.30)	(0.38)	(0.00)
Tobin's q			0.08754***
			(7.22)
Cash to Assets			0.46847***
			(5.60)
Dividend Yield			-0.08163***
			(-20.66)
Historical Volatility			0.86861***
			(-19.44)
Leverage			-0.00184***
			(-5.17)
Size			0.0000***
			(2.80)
Fixed Effects	No	Yes	Yes
R^2	0.0013	0.0012	0.2521

This table reports the results of regressions of the excess returns on firms' separate environmental, social, and governance score under several specifications: without firm controls (specification 1), with fixed effects (specification 2), and with fixed effects and firms controls (specification 3). Control variables are winsorized at the 1% level in each tail (except for Size at the 5% level). Standard errors are heteroscedasticity robust. The regression constant is not reported for brevity. The t-statistics are reported in parentheses. Table A1 in the appendix defines all the variables used. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

When diving the total effect of Regression (1) into its three pillars, the separate effects can be measured. Again, the first specification shows the results without fixed effects, nor control variables. However, these coefficients could be biased, since they do not account for the correlation between the error term and the predictor variables. When adding fixed effects, but no control variables, both the environmental and the social pillar are significant, with respectively a negative coefficient of -0.15% and a positive coefficient of 0.11%.

Since the R^2 of the model including the control variables is much higher (0.25 compared to 0.001) for the third specification, this is again the most appropriate model to measure the relation. After adding the control variables, only the environmental pillar remains significant at 1%, with a coefficient of -0.00095. In other words, when the environmental score of a firm increases with 1 point, the yearly excess returns will decrease with on average 0.1%. Hence, the observed effect from Regression (1) of 0.12% is nearly completely driven by the environmental score, and not by the social, nor governance score.

In conclusion, the first hypothesis that states that there is a positive relation between ESG scores and performance does not hold for this sample. There is instead a significant, negative effect of ESG scores on yearly excess returns, which is strongly driven by the environmental score.

4.2 The Performance of High ESG Rated Firms during the Pandemic

Regarding the second hypothesis, which looks at the relation between ESG scores and performance during the COVID-19 period, the Regressions (3), (4), and (5) will be used. The results are shown in Table 4.

Table 4. Panel regression of yearly excess returns on high ESG rated firms during Covid-19

This table shows the effect of having a high ESG score during the pandemic on firm performance including firm controls. No overall effect is found. When splitting this effect in two years, there is also no effect present during 2019, but a significant positive effect during 2020.

	(1)	(2)	(3)
Dependent variable	Excess return	Excess return	Excess return
High ESG Score	-0.05483***	-0.04451***	-0.02741***
	(-5.27)	(-4.31)	(-2.71)
Covid Period	0.15351***		
	(9.14)		
High ESG * Covid Period	0.01962		
	(1.02)		
Covid Year 1		0.15647***	
		(9.51)	
High ESG * Covid Year 1		0.01657	
		(0.86)	
Covid Year 2			-0.03234
			(-1.20)
High ESG * Covid Year 2			0.10686***
			(3.49)
Tobin's q	.07431***	0.07850***	0.07918***
	(7.00)	(7.37)	(7.37)
Cash to Assets	0.30636***	0.32667***	0.46518***
	(3.83)	(4.04)	(5.69)
Dividend Yield	-0.07743***	-0.08033***	-0.07811***
	(-20.48)	(-20.61)	(-20.06)
Historical Volatility	-1.03685***	-0.90994***	-0.99631***
	(-23.96)	(-20.19)	(-20.90)
Leverage	-0.00243***	-0.00239***	-0.00176***
	(-7.28)	(-7.13)	(-5.37)
Size	0.00000**	0.00000***	0.00000***
	(1.98)	(2.71)	(2.92)
Fixed effects	Yes	Yes	Yes
R^2	0.2669	0.2643	0.2526

This table reports the results of regressions of the excess returns on firms' ESG score and a dummy variable representing the Covid-19 period, under several specifications: the entire Covid-period (specification 1), the first year of Covid (specification 2), and the last year of Covid (specification 3). Control variables are winsorized at the 1% level in each tail (except for Size at the 5% level). Standard errors are heteroscedasticity robust. The regression constant is not reported for brevity. The t-statistics are reported in parentheses. Table A1 in the appendix defines all the variables used. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

The first specification reports the regression of the excess returns on a dummy variable for high ESG rated firms, a dummy variable representing the entire pandemic period up until the end of 2021, and an interaction effect of the two. The latter is the variable that will answer the second hypothesis. Because of the interaction terms, the interpretation of all the betas will change in the model, which is why only the sign and significance can be assessed. Fixed effects are added, since the Hausman test shows that the error terms and the predictor variables are correlated.

The coefficient for the interaction effect between high ESG rated firms and the Covid period under the first specification appears to be nonsignificant, meaning that there is no observed effect for firms being in the high ESG rated group during Covid time, relative to the low ESG group. In other words, there is no significant difference in returns because of the pandemic for high ESG rated firms compared to low ESG rated firms. When splitting the two years captured by the Covid Period variable in specification 2 and 3, it can be seen that the interaction effect between high rated firms and the first year of Covid is insignificant, while in 2021, on the other hand, the interaction effect does appear to be significant and positive. This means that for the last year of the sample, being a high ESG rated firm during 2021 has a positive effect on the yearly excess returns compared to being a low ESG rated firm. All three specifications have a comparable R^2 around 0.262.

In conclusion, the second hypothesis that states that high ESG rated stocks have outperformed low ESG rated stocks during Covid-19, is partially rejected. The overall interaction effect is insignificant for this sample, meaning that there has been no significant difference between high and low ESG rated stocks during the entire pandemic in the sample. However, when splitting this effect for the two separate years, during the second year of Covid, High ESG rated firms did have higher excess returns relative to low rated firms.

4.3 Are Upgrades Rewarded more than Downgrades are Being Punished?

Next, the third hypothesis is being tested, which states that ESG score upgrades are rewarded more than ESG downgrades are being punished. Regressions (6) and (7) are used, and the results are presented in Table 5 and Table 6.

Table 5. Panel regression of yearly excess returns on ESG score upgrades versus score downgrades

This table shows the effect of an ESG upgrade relative to a downgrade on firm performance including firm controls. The ESG upgrade variable is significant and negatively related to performance.

	(1)	(2)
Dependent variable	Excess return	Excess return
ESG score upgrade	-0.02717***	-0.02893***
	(-3.26)	(-4.03)
Tobin's q		0.08072***
		(11.47)
Cash to Assets		0.36790***
		(6.88)
Dividend Yield		-0.07854***
		(-29.85)
Historical Volatility		-0.87865***
		(-23.93)
Leverage		-0.00206***
		(-8.38)
Size		0.0000*
		(1.66)
Fixed effects	Yes	Yes
R^2	0.0002	0.2230

This table reports the results of regressions of the excess returns on a dummy variable containing ESG upgrades under two specifications: without firm controls (specification 1), and with firm controls (specification 2). Control variables are winsorized at the 1% level in each tail (except for Size at the 5% level). Standard errors are heteroscedasticity robust. The regression constant is not reported for brevity. The t-statistics are reported in parentheses. Table A1 in the appendix defines all the variables used. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Under the first specification, the simple regression of yearly excess returns on the dummy variable ESG score upgrade is shown. The coefficient of -0.0272 is significant at the 1% level, meaning that when the change in a firm's ESG score is positive, the excess returns are about 2.7% lower than for a negative change in ESG score. This finding also holds when adding the control variables in the second specification, where the excess returns are around 2.9% lower for an upgrade compared to a downgrade. This means that score upgrades are being punished instead of rewarded, relative to a score downgrade, which points into the opposite direction of the third hypothesis.

The R^2 increases when adding control variables from 0.0002 to 0.22, which makes the second one the most fitting model.

Table 6. Panel regression of yearly excess returns on ESG Momentum group variables

This table shows the effect of 9 different momentum groups on firm performance including firm controls, with group 3 being the reference group containing small to no change in rating. There is no effect of ESG downgrades (group 1 and 2) on performance, and a significant negative effect for ESG upgrades higher than 20.5 points. The negative effect increases as the score upgrade becomes higher.

	(1)	(2)
Dependent variable	Excess return	Excess return
Group 1	-0.03294	0.01657
	(-1.06)	(0.72)
Group 2	-0.00178	0.01093
	(-0.11)	(0.68)
Group 3	*Reference group*	*Reference group*
Group 4	-0.01943	-0.00724
	(-1.21)	(-0.47)
Group 5	-0.03066	-0.02254
	(-1.61)	(-1.23)
Group 6	-0.05130*	-0.00444*
	(-1.90)	(-1.86)
Group 7	-0.10897***	-0.08094***
	(-3.58)	(-2.88)
Group 8	-0.13253***	-0.11421***
	(-3.62)	(-3.35)
Group 9	-0.13090***	-0.07285**
	(-4.06)	(-2.44)
Tobin's q		0.08256***
		(7.39)
Cash to Assets		0.45873***
		(5.53)
Dividend Yield		-0.08071***
		(-20.69)
Historical Volatility		-0.85040***
		(-17.83)
Leverage		-0.00167***
		(-4.80)
Size		0.0000
		(1.17)
Fixed effects	Yes	Yes
R^2	0.0032	0.2517

This table reports the results of regressions of the excess returns on 9 dummy variables containing nine different groups of ESG Momentum values, ranging from -16 until 60, under two specifications: without firm controls (specification 1), and with firm controls (specification 2). Control variables are winsorized at the 1% level in each tail (except for Size at the 5% level). Standard errors are heteroscedasticity robust. The regression constant is not reported for brevity. The t-statistics are reported in parentheses. Table A1 in the appendix defines all the variables used. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 6 reports the results of the regression of excess returns on the different dummy variables containing low to high ESG momentum values. The third group is omitted because of multicollinearity, and since it contains values around zero, it serves as the reference group. The R^2 of the model where controls are included is 0.25 in contrast to 0.003 without controls, but the coefficients are roughly the same for both models. The groups containing the downgrades, group 1 and 2, are both highly insignificant. This means that no effect is measured of neither low, nor high downgrades on the excess returns. When moving up from the reference group, the t-statistics start increasing, and the sixth group, containing ESG momentum values between 20.5 and 30.5, is the first variable to be significant at the 10% level, with a coefficient of -0.004. For groups 7, 8 and 9, the coefficients are all negative and significant at 1%, with the largest effect for positive score changes between 40.5 and 50.5, of -0.1142.

It can thus be concluded that the higher the ESG score upgrade, the bigger the punishment in terms of returns. Regarding downgrades, no effect on the excess returns is observed, hence, it is neither being punished, nor rewarded. This means that the third hypothesis is rejected, since ESG score upgrades are punished more than ESG score downgrades, instead of the expected other way around.

4.4 Is ESG Momentum a Better Predictor?

To test the fourth and last hypothesis which states that ESG momentum is a better predictor for performance, relative to standalone ESG scores, Regression (8) will be used. The results are presented in Table 7.

Table 7. Panel regression of yearly excess returns on the change in ESG scores

This table shows the effect of ESG momentum on firm performance including firm controls. The ESG momentum variable is highly significant and negatively related to firm performance. When including both standalone and momentum scores, both variables are still significant and negatively related to performance.

	(1)	(2)	(3)	(4)
Dependent variable	Excess return	Excess return	Excess return	Excess return
ESG Momentum	-0.00145***	-0.00209***	-0.00182***	-0.00191***
	(-5.59)	(-6.54)	(-6.71)	(-6.78)
ESG Score				-0.00137***
				(-4.15)
Tobin's q			0.07935***	0.08079***
			(7.43)	(7.07)
Cash to Assets			0.44816***	0.48414***
			(5.38)	(5.63)
Dividend Yield			-0.0803***	-0.08257***
			(-20.80)	(-19.98)
Historical Volatility			-0.85846***	-0.88348***
			(-18.50)	(-17.93)
Leverage			-0.00165***	-0.00163***
			(-4.88)	(-4.66)
Size			0.00000	0.00000***
			(1.43)	(2.63)
Fixed effects	No	Yes	Yes	Yes
R^2	0.0027	0.0027	0.2510	0.2660

This table reports the results of regressions of the excess returns on firms' change in ESG score under several specifications: without firm controls (specification 1), with fixed effects (specification 2), with fixed effects and firms controls (specification 3), and with ESG Score as added variable (specification 4). Control variables are winsorized at the 1% level in each tail (except for Size at the 5% level). Standard errors are heteroscedasticity robust. The regression constant is not reported for brevity. The t-statistics are reported in parentheses. Table A1 in the appendix defines all the variables used. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

The first specification shows the regression of the yearly excess returns on the year-on-year change in ESG ratings, with a one year lag. The coefficient of -0.0015 is highly significant. For the ESG momentum variable, just as for the regular ESG scores, the Hausman test shows that the error terms and the predictor variables are correlated, which is why in the second specifications, fixed effects are included into the model.

The coefficient than changes to -0.0021, meaning the effect has become more negative, while still significant at the 1% level. When the control variables are added, the R^2 of the model increases from 0.003 to 0.25, hence, the coefficient of ESG momentum under the third specification will be the most unbiased of the three. The coefficient of -0.00182 is also significant at the 1% level, meaning that when the year-on-year change in ESG rating goes up with one point, the yearly excess returns on average lower with 0.18%. When comparing this to the coefficient of the ESG Score (Regression 1, Table 2) of -0.12%, it can be concluded that the negative effect on the yearly excess returns is greater for a change in the ESG ratings, than for standalone ESG scores. This is confirmed when adding ESG score to the model, under specification four. ESG momentum has a significant coefficient of -0.0019, while ESG Score has one of -0.0014, hence, the negative effect is indeed larger for ESG momentum than for the standalone ESG score. However, comparing the betas cannot assess whether the ESG momentum variable is a better predictor for returns than the ESG rating, since both variables have a different range (the score from 0 to 100, and momentum from -16 to 60), hence, both betas have a different scale. A more suitable way to test the fourth hypothesis, is to compare the R^2 of the two models. The first one explains 25.19% of the variation in the model, while the last one explains 25.12%, meaning both models roughly do an equally good job at fitting the data, and ESG momentum turns out not to be a better predictor for yearly excess returns than ESG score. When adding both variables to the model, the R^2 is 0.2660, which is somewhat higher, but a negligible difference.

In conclusion, both a higher ESG momentum and a higher standalone ESG score lead to a lower yearly excess return, with respectively 0.19% and 0.14% per upgraded point, both significant at the 1% level. When comparing both individual models, both R^2 are approximately the same, and it does not increase much when combining the two models. This means that the fourth hypothesis is rejected, since the ESG momentum variable is not a better predictor for performance than the standalone ESG rating.

5. Discussion

According to the above presented results, the first hypothesis that states that there is a positive relationship between ESG scores and performance cannot be supported. The relationship shows to be significantly negative; when the ESG score increases with one point, the yearly excess returns lower by 0.12%. This result is almost completely driven by the environmental pillar, leaving the social and governance score to be irrelevant. This effect could possibly be explained by the carbon premium, proposed by Bolton and Kacperczyk (2020), meaning that more polluting firms yield higher returns while less polluting firms (high E-score firms) show lower returns. The overall results point towards the earlier mentioned sin stock hypothesis, which states that sin stocks tend to outperform their benchmarks, because they are undervalued (Merton, 1987). The results are also in line with earlier research that has found that investors are often willing to pay a premium for making sustainable investment decisions, leading to the overpricing of high ESG rated firms (Zhong & Chen, 2019). Another explanation is provided by Pastor et al. (2020); high ESG-rated stocks are less risky since they function as a hedge, thus investors like holding these stocks, leading to lower returns.

The second hypothesis stating that high ESG rated stocks have outperformed low ESG rated stocks during the pandemic is partially rejected. The overall interaction effect between high ESG rated stocks and the pandemic period is insignificant, yet when looking only at 2020, there is a positive effect present. A possible explanation for the different effect in the two years is that during the first year, people were facing a lot of uncertainty which might make them choose the "safer" investment options that are always needed such as oil, gas, and gold (Rubbaniy et al., 2021). In contrast to the perspective of Lööf et al. (2022), the findings in this research do not point towards sustainable firms being considered safer during uncertainty. In the second year of the pandemic people gained more perspective, a vaccine was developed and things were looking up again, which might have led to investors choosing more sustainable options, resulting in a positive return for high ESG stocks.

The overall insignificant effect of sustainable stocks during the pandemic is consistent with the findings of Bae et al. (2021), and contradict the findings of Broadstock et al. (2022) and Albuquerque et al. (2020). The contradiction can arise because of the use of different datasets; the first mentioned research looks solely at China, and the latter uses only the first quarter of 2020.

The third hypothesis stated that ESG score upgrades are rewarded more than downgrades are being punished. Based on the results, this hypothesis is rejected, since the opposite was found. The higher the ESG score upgrade, the bigger the "punishment" in terms of return. The returns are about 2.9% lower for an ESG upgrade compared to a downgrade. When specifying the results further, it was found that there was no effect for ESG downgrades on returns, and ESG upgrades from 20.5 and higher were being punished. The largest punishment was found for an upgrade between 40.5 and 50.5 of -0.11. These finding contradict the findings of Giese and Nagy (2018), who used the MSCI ESG scores. The difference in findings shows the possible effect of the drastic divergence in ratings between scoring companies, since the MSCI and Asset4 database only show a correlation between ESG scores of 38% (Berg et al., 2019). Furthermore, given the findings from the first hypothesis, the results are in line with expectations; if a higher standalone ESG score leads to lower returns, it makes sense that a positive change in ESG scores also yields lower returns. Lastly, whether the findings are in line with Prospect Theory (Kahneman and Tversky, 1979), depends on whether score upgrades are being regarded as positive or negative by investors. In the case investors see an upgrade as a positive event and a downgrade as a negative one, more weight is being put on the positive event, which would contradict Prospect Theory. However, since the effect of a score upgrade is negative, it is plausible that a score upgrade is actually regarded as a negative event, and a downgrade as a positive one. In this case, the results would reflect loss aversion, which would be consistent with Prospect Theory.

When looking at ESG momentum as a predictor for returns, the results show that it does not do a better job than the standalone ESG rating, which leads to a rejection of the fourth hypothesis. Both ESG momentum and ESG score lead to a lower yearly excess return with respectively 0.19% and 0.14% per upgraded point. Since both variables have a completely different scale, the betas cannot be compared to assess which variable is a better predictor.

When looking at the R^2 of both models, the difference is neglectable, meaning that one is not a better predictor than the other. It is noteworthy that when including both ESG momentum and ESG score in the model, both variables are highly significant, meaning that they likely capture something different. The results contradict the earlier literature on ESG momentum, for example by Khan et al. (2016), Nagy et al. (2016), and Giese et al. (2019). This again could be attributed to the use of different ESG ratings such as MSCI, and the use of pure returns instead of excess returns.

Combining all results, the main question can be answered: *What is the effect of ESG activities on a firm's financial performance and does it offer a safe haven during a crisis period?*

There is an overall significant negative effect of ESG activities on a firm's financial performance, reflected through both ESG ratings and the year-on-year change in ESG ratings. This effect is mainly driven by the environmental score. The higher the increase in ESG score, the larger the subsequent "punishment" in terms of excess return. High ESG rated stocks have not offered a complete safe haven during the pandemic, but only partially in the second year. Thus, investing in high ESG rated stocks might not be the best decision when one is only looking to maximize returns. However, when one is looking to contribute to a sustainable and social world, and is willing to pay some premium to contribute, investing in sustainable firms might not be such a bad decision after all.

6. Limitations and Future Research

In the following section, the limitations of the data, methodology and generalizability of this paper will be discussed. Firstly, the used sample consists of all European firms listed on Asset4. Therefore, the observed results are only applicable to the European market and may possibly not hold for other countries.

Another limitation is formed by the measurement of how ESG scores and score changes are interpreted by investors. Stock prices are used to measure how investors response to a particular (change in) ESG score, while this may not fully capture their attitude. An increase in stock prices because of increased attention of investors do not form the best indicator for long-term value creation, but rather capture short-term effect. The reported relationships are therefore of a more short-lived nature, instead of capturing long-term value creation as a result of increased sustainability within firms.

Furthermore, another limitation is consequence of the divergence in ESG ratings between different scoring companies. This study focusses only on the Asset4 database from Refinitiv, but because of the low correlation between the scoring firms, it might be that different results would have been reported using another database.

Next, this research uses the yearly excess stock returns as a proxy for performance, which forms a limitation for two reasons. Firstly, by looking at yearly stock returns, a lot of valuable information and developments throughout the year are lost. It forms a rather simplified image of reality. Secondly, no formal European risk-free rate was used during the calculations for the simple reason that it does not exist yet. The 3 month EURIBOR rate is used as a proxy for the European risk-free rate, which may have altered the results.

Another limitation is the assumption that ESG scores are updated at the beginning of the year, so that the effect could be measured over that particular year. This is however not the case in reality since the ESG data disclosure is unstructured (Refinitiv ESG Scores Factsheet, 2022).

For the last section, subsequent areas of possible future research are identified. Firstly, it would be interesting to look into the differences between different geographies. According to Kaiser (2020), the adoption rate of ESG is higher in Europe, which could lead to a more efficient way of pricing ESG scores (changes) compared to other countries.

Secondly, it could be valuable to take a more long-term approach and use another method to measure valuation by investors. Long-term value creation as a result of sustainable returns could then be measured, which would provide another perspective on the relationship between sustainability and performance.

Next, the same research design could be used for different scoring companies. This way it could be tested to what extent the divergence in ratings would lead to a divergence in results. This is important to establish, since using merely one database could distort the reported relationships, leading to misleading practical implications.

Another research implication is to measure monthly effects instead of yearly effects. It would be interesting to study possible patterns after an ESG score change in the following period, instead of looking just at the first month of the next year. Moreover, as ESG scores become increasingly important, they might be updated on a monthly basis in the future, which would again allow to identify patterns as well as immediate effects.

It would also be relevant to take the actual ESG score update into account, instead of assuming it happens at the beginning of the year. This would allow for a more precise and accurate research, with possibly different results.

Lastly, since the results have shown that when including both ESG momentum and ESG score in the model as predictors, both variables are highly significant, which means that they capture something different. It would be interesting to dive further into these effects and identify what it is they capture.

Appendix

Table 1A. Variables, definitions, and sources

Variable	Definition	Source
Yearly excess return	Using the unpadded adjusted prices, the yearly excess return is calculated as $\ln(P_{t+1} / P_t)$, in excess of the 3-month EURIBOR rate	Prices: Datastream EURIBOR: Euribor-rates.eu
ESG rating	Refinitiv's ESG Score is an overall company score based on the self-reported information in the environmental, social and corporate governance pillars.	Thomson Reuter's Refinitiv ESG
E score	Refinitiv's Environment Pillar Score is the weighted average relative rating of a company based on the reported environmental information and the resulting three environmental category scores.	Thomson Reuter's Refinitiv ESG
S score	Refinitiv's Social Pillar Score is the weighted average relative rating of a company based on the reported social information and the resulting four social category scores.	Thomson Reuter's Refinitiv ESG
G score	Refinitiv's Governance Pillar Score is the weighted average relative rating of a company based on the reported governance information and the resulting three governance category scores.	Thomson Reuter's Refinitiv ESG
$HighESG_t$*$CovidPeriod_t$	Interaction effect between a dummy variable equal to one when the ESG rating is in the top 50% of scores, and a dummy variable equal to one during the pandemic years in the sample, 2020 and 2021.	
$HighESG_t$*CovidYear1	Interaction effect between a dummy variable equal to one when the ESG rating is in the	

	top 50% of scores, and a dummy variable equal to one during the first pandemic year in the sample, 2020.	
$HighESG_t$*CovidYear2	Interaction effect between a dummy variable equal to one when the ESG rating is in the top 50% of scores, and a dummy variable equal to one during the last pandemic year in the sample, 2021.	
$Upgrades_t$	Dummy variable with value one when ESG Momentum is positive, and zero otherwise	
$G1_t$ – $G9_t$	Dummy variable groups where $G1_t$ captures the lowest ESG Momentum values, and $G9_t$ the highest ones at time *t.*	
ESG $Momentum_t$	the year-on-year change in ESG ratings, accounting for a one year time lag.	
Tobin's q	Market value of equity plus book value of equity of preferred stock and debt divided by book value of total assets.	Thomson Reuter's Datastream
Size	The share price multiplied by the number of ordinary shares in issue, expressed in millions of units of local currency.	Thomson Reuter's Datastream
Cash to assets	Cash holdings over total assets	Thomson Reuter's Datastream
Leverage	(Long Term Debt + Short Term Debt & Current Portion of Long Term Debt) / (Total Capital + Short Term Debt & Current Portion of Long Term Debt) * 100	Thomson Reuter's Datastream
Historical volatility	Volatility of a stock over the time *t-1* until *t.*	Thomson Reuter's Datastream
Dividend yield	The dividend per share as a percentage of the share price. Underlying dividend is based on an anticipated annual dividend and excludes special or once-off dividends.	Thomson Reuter's Datastream

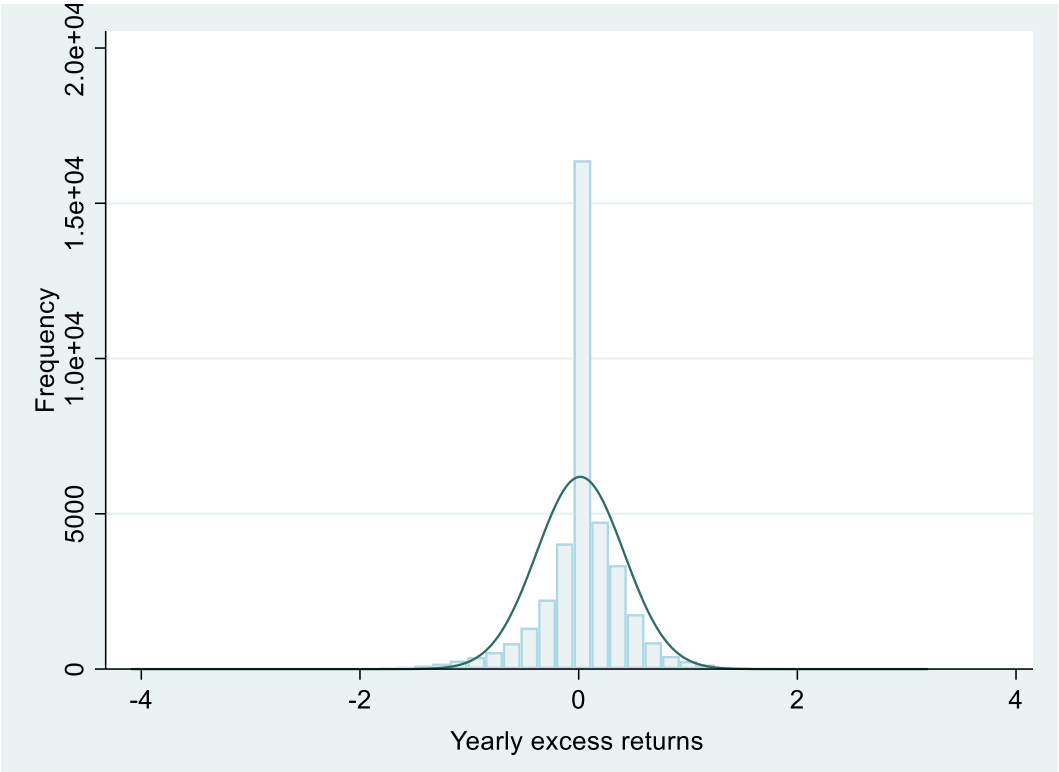

Figure 5. Distribution of yearly excess returns

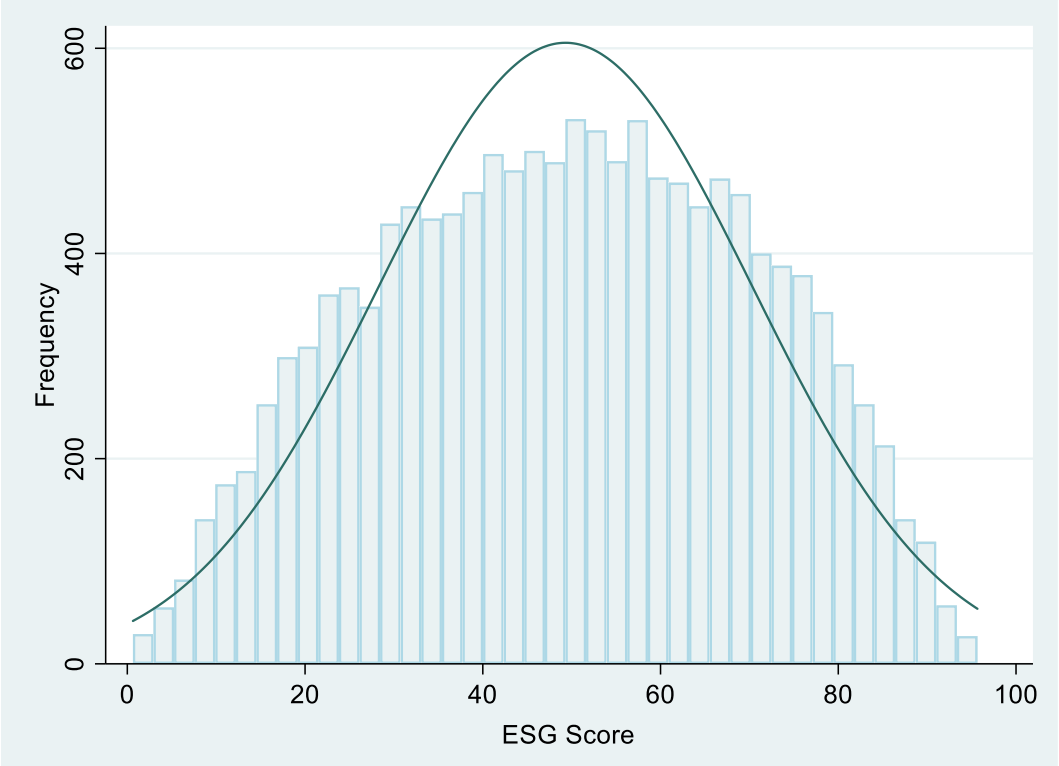

Figure 6. Distribution of ESG scores

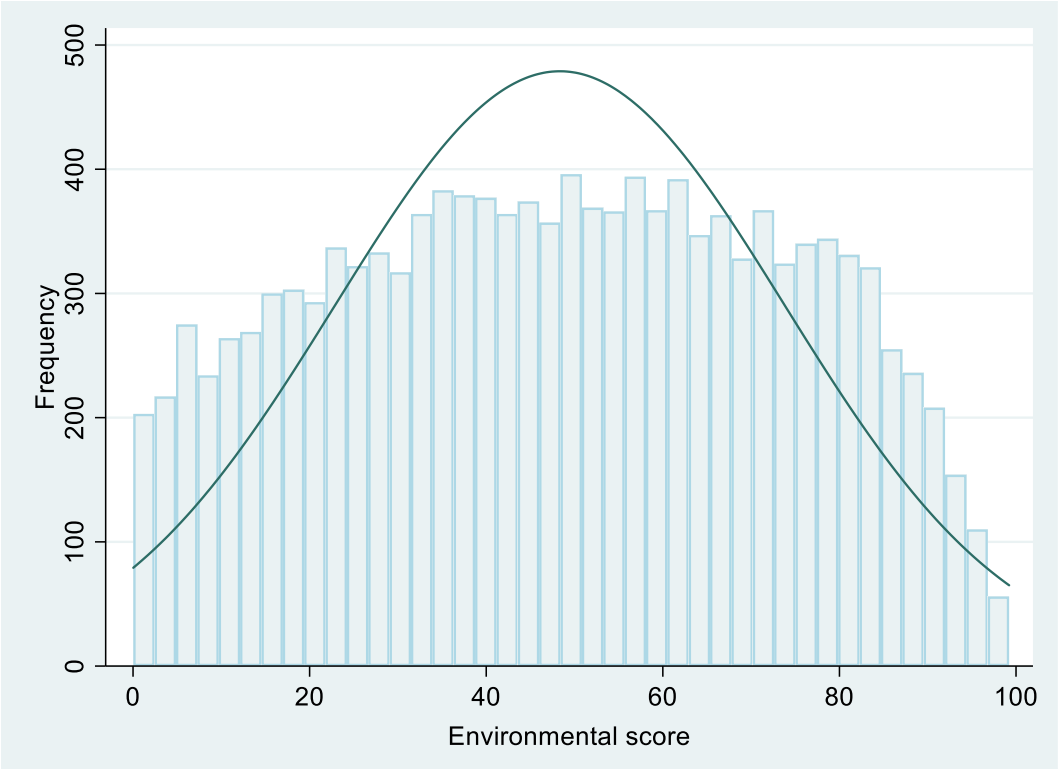

Figure 7. Distribution of E score

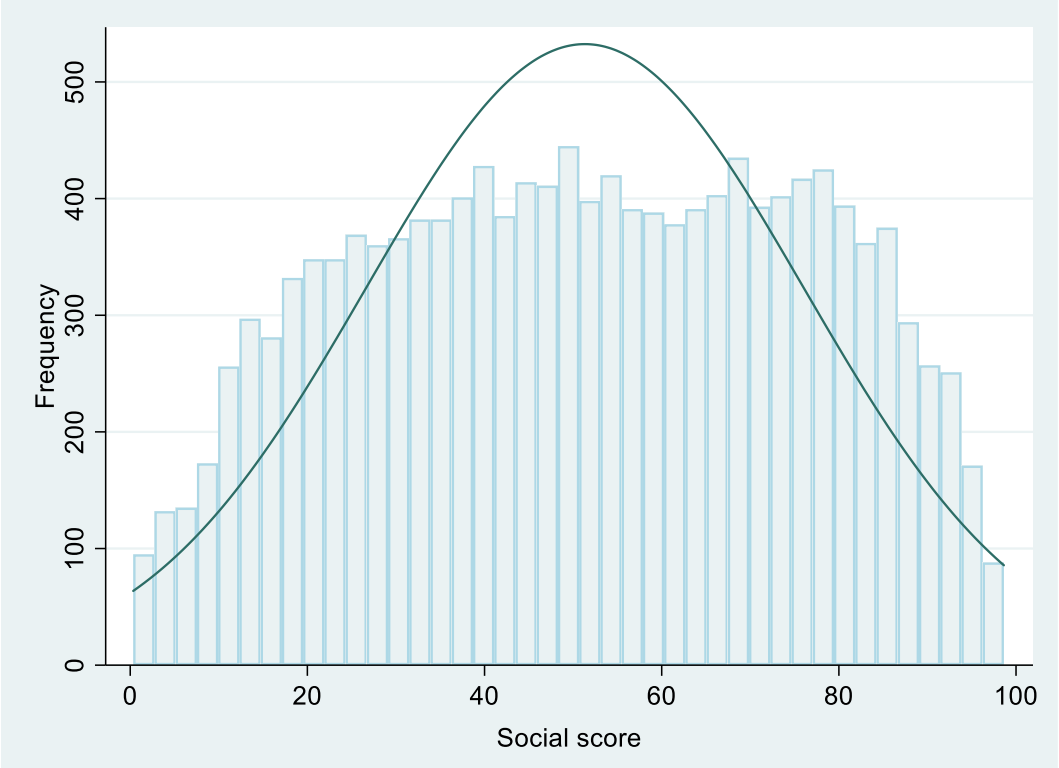

Figure 8. Distribution of S score

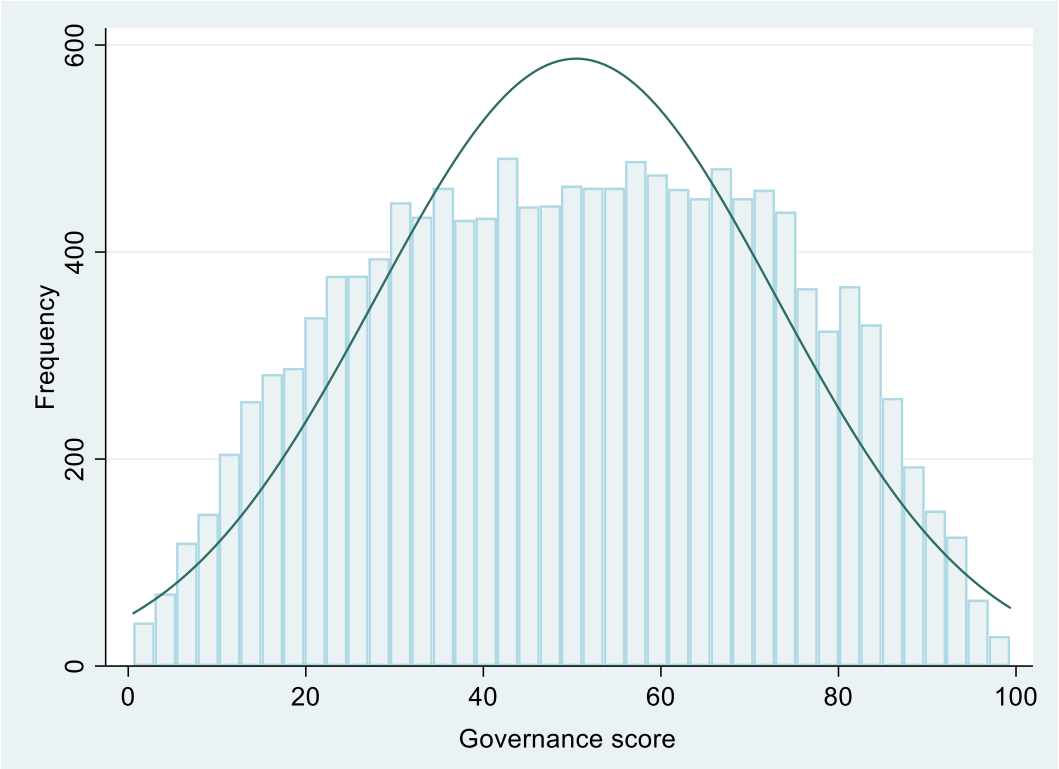

Figure 9. Distribution of G score

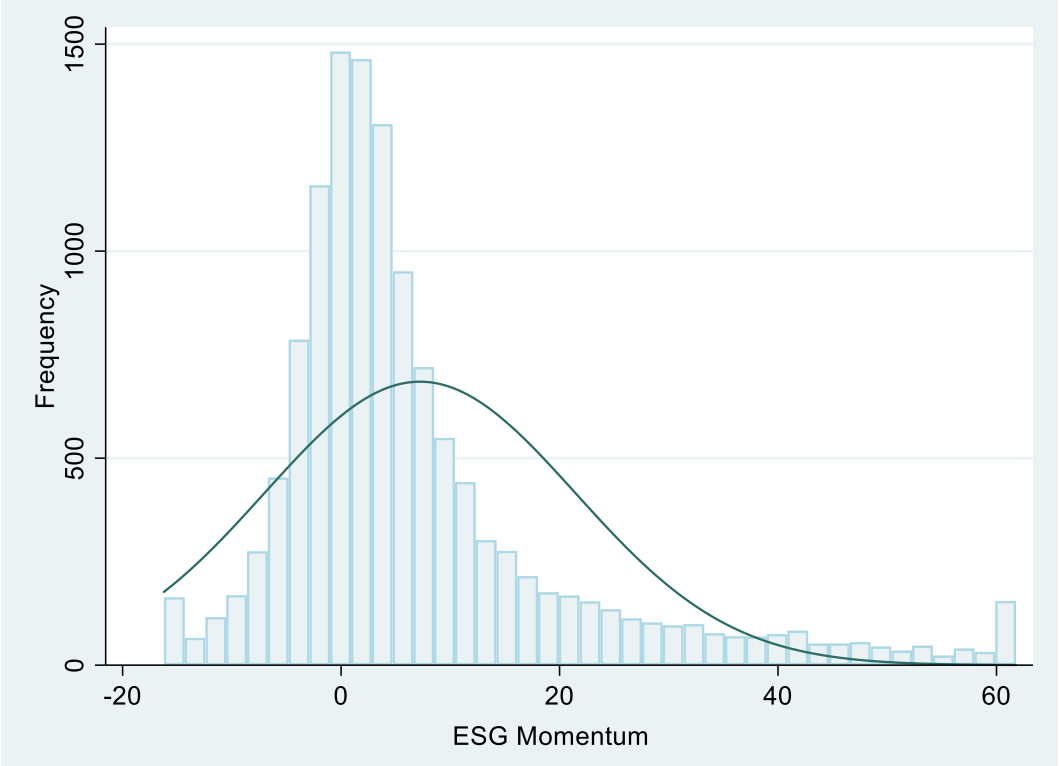

Figure 10. Distribution of ESG Momentum

Table 2A. Summary statistics of ESG Momentum per group.

ESG Momentum by group	Observations	Mean	Std. dev.	Minimum	Maximum
Group 1	348	-13.91	2.16	-16.22	-10.51
Group 2	3,183	-3.69	2.45	-10.49	-0.51
Group 3	814	-0.01	0.29	-0.50	0.49
Group 4	5,339	4.48	2.71	0.50	10.49
Group 5	1,437	14.56	2.82	10.50	20.49
Group 6	680	24.96	2.86	20.50	30.47
Group 7	416	35.19	2.92	30.50	40.49
Group 8	304	45.03	2.92	40.50	50.47
Group 9	330	58.09	3.84	50.50	61.79

References

Adams, C. A. (2017). *Understanding integrated reporting: The concise guide to integrated thinking and the future of corporate reporting*. Routledge.

Albuquerque, R., Koskinen, Y., Yang, S., & Zhang, C. (2020). Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. *The Review of Corporate Finance Studies*, *9*(3), 593-621.

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, *74*(3), 87-103.

Barber, B. M., Morse, A., Yasuda, A. (2021). Impact investing. *Journal of Financial Economics, 139(1)*, 162–185. https://doi.org/10.1016/j.jfineco.2020.07.008

Bennani, L., Le Guenedal, T., Lepetit, F., Ly, L., Mortier, V., Roncalli, T., & Sekine, T. (2018). How ESG Investing has impacted the asset pricing in the equity market. *Available at SSRN 3316862*.

Berg, F., Koelbel, J. F., & Rigobon, R. (2019). *Aggregate confusion: The divergence of ESG ratings* (pp. 1-42). Cambridge, MA, USA: MIT Sloan School of Management.

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG Ratings: (Dis) agreement and performance. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1426-1445.

Bloomberg Intelligence (2021). ESG assets may hit $53 trillion by 2025, a third of global AUM. Retrieved from https://www.bloomberg.com/professional/blog/esg-assets-may-hit-53-trillion-by-2025-a-third-of-global-aum/

Bolton, P., & Kacperczyk, M. T. (2020). Carbon premium around the world.

Bouten, L., Cho, C. H., Michelon, G., & Roberts, R. W. (2017). CSR Performance Proxies in Large-Sample Studies:'Umbrella Advocates', Construct Clarity and the 'Validity Police'. *Construct Clarity and the 'Validity Police' (August 2017).*

Brown, D. L., Guidry, R. P., & Patten, D. M. (2009). Sustainability reporting and perceptions of corporate reputation: An analysis using fortune. In *Sustainability, environmental performance and disclosures*. Emerald Group Publishing Limited.

Buallay, A. (2018). Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector. *Management of Environmental Quality: An International Journal*.

Cheng, B., Ioannou, I., & Serafeim, G. (2014). Corporate social responsibility and access to finance. *Strategic management journal*, *35*(1), 1-23.

Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. *Journal of Banking & Finance*, *35*(8), 2137-2147.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The accounting review*, *86*(1), 59-100.

Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, *167*(2), 333-360.

Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, *131*(3), 693-714.

ECB (2022). Working group on euro risk-free rates. Retrieved from: https://www.ecb.europa.eu/paym/interest_rate_benchmarks/WG_euro_risk-free_rates/html/index.en.html

Edelman (2019). An overwhelming majority of investors expect companies to implement effective environmental, social-impact and governance (ESG) practices. Retrieved from: https://www.edelman.com/news-awards/2019-institutional-investor-trust.

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics, 101(3),* 621-640.

Edwards, K. D. (1996). Prospect theory: A literature review. *International review of financial analysis*, *5*(1), 19-38.

Environmental, Social, and Governance Scores Methodology. *Refinitiv,* March 2022.

Ferrell, A., Liang, H., & Renneboog, L. (2016). Socially responsible firms. *Journal of financial economics*, *122*(3), 585-606.

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.

Friedman, M. (2007). The social responsibility of business is to increase its profits. In *Corporate ethics and corporate governance* (pp. 173-178). Springer, Berlin, Heidelberg.

Fulton, M., Kahn, B., & Sharples, C. (2012). Sustainable investing: Establishing long-term value and performance. *Available at SSRN 2222740*.

Giese, G., Lee, L. E., Melas, D., Nagy, Z., & Nishikawa, L. (2019). Foundations of ESG investing: How ESG affects equity valuation, risk, and performance. *The Journal of Portfolio Management*, *45*(5), 69-83.

Giese, G., & Nagy, Z. (2018). How markets price ESG. *Have Changes in ESG Scores Affected Stock Prices*, 1-25.

Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility—Where are they? A critical review of ESG investing. *Review of Financial Economics*, *26*, 25-35.

Hanicova, D., & Vojtko, R. (2020). Backtesting ESG Factor Investing Strategies. *Available at SSRN 3633303*.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics*, *93*(1), 15-36.

Hvidkjær, S. (2017). ESG investing: a literature review. *Report prepared for Dansif*.

Kahneman, D. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47*(2), 264-291.

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The accounting review*, *91*(6), 1697-1724.

Kishan, S. (2022). ESG by the Numbers: Sustainable Investing Set Records in 2021. *Bloomberg.*

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *the Journal of Finance*, *72*(4), 1785-1824.

Lioui, A. (2018). Is ESG Risk Priced? (SSRN Scholarly Paper ID 3285091). Social Science Research Network. https://doi.org/10.2139/ssrn.3285091

Lööf, H., Sahamkhadam, M., & Stephan, A. (2022). Is Corporate Social Responsibility investing a free lunch? The relationship between ESG, tail risk, and upside potential of stocks before and during the COVID-19 crisis. *Finance Research Letters*, *46*, 102499.

Nagy, Z., Kassam, A., & Lee, L. E. (2016). Can ESG add alpha? An analysis of ESG tilt and momentum strategies. *The Journal of Investing*, *25*(2), 113-124.

Paribas, B. N. P. (2017). Great Expectations For ESG. What's next for Asset Owners and Managers.

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2020). Sustainable investing in equilibrium. *Journal of Financial Economics.*

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, *142*(2), 572-597.

Principles for Responsible Investment Brochure 2021 (English), from https://www.unpri.org/about-us/about-the-pri

Refinitiv Eikon (2022). ESG Scores Fact sheet, from https://www.refinitiv.com/content/dam/marketing/en_us/documents/fact-sheets/esg-scores-fact-sheet.pdf

Renneboog, L., Ter Horst, J., and Zhang, C. (2008b). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of corporate finance, 14*(3), 302-322.

Rubbaniy, G., Khalid, A. A., Rizwan, M. F., & Ali, S. (2021). Are ESG stocks safe-haven during COVID-19?. *Studies in Economics and Finance*.

Simnett, R., Vanstraelen, A., & Chua, W. F. (2009). Assurance on sustainability reports: An international comparison. *The accounting review*, *84*(3), 937-967.

Statman, M., & Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, *65*(4), 33-46.

The Materiality of Social, Environmental and Corporate Governance Issues to Equity Pricing, *UNEP FI,* June 2004.

Torres-Reyna, O. (2007). Panel Data Analysis; Fixed and Random Effects using Stata, Princeton University.

Von Neumann, J., & Morgenstem, 0. (1953) *Theory of Games and Economic Behavior*, Princeton University Press. Princeton, New Jersey.

Weber, E. U., & Johnson, E. J. (2009). Decisions under uncertainty: Psychological, economic, and neuroeconomic explanations of risk preference. In *Neuroeconomics* (pp. 127-144). Academic Press.

Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance*, *98*, 39-60.

Zhong, S., & Chen, J. (2019). How environmental beliefs affect consumer willingness to pay for the greenness premium of low-carbon agricultural products in China: theoretical model and survey-based evidence. *Sustainability*, *11*(3), 592.