Expanding the Fama & French Models with ESG scores: Evidence from the Russell 1000 index

Name: Jonathan Dol

Student number: 550653

Email: 550653jd@student.eur.nl

Erasmus Universiteit Rotterdam,

Economie en Bedrijfseconomie

Research question: What is the effect of ESG scores on excess stock returns?

Supervisor: Esad Smajlbegovic

Co-reader: Giovanni Cocco

Thesis language: English

Abstract

This research paper explores the possibilities of expanding the works of the Fama & French Three-Factor and Fama & French Five-Factor Models (1992, 2015). This paper tests several hypotheses to expand them with not only the ESG score, but to also look at the possibilities to expand the Fama & French models with the Environment-, Social-, or Governance score. For this research, the paper looks at the companies from the Russel 1000, during the period of 2010 till 2020. This paper found the following results.

Looking at the results of the Fama & MacBeth (1973) regression with the Three-Factor model control variables, it is found that hypotheses 1, 2, and 3 are all rejected at the 5% significance level. For the fourth hypothesis, it is found that the Governance score coefficient is statistically significant. It is found that on average, in our sample, when the Governance score is one point higher, the average monthly stock return becomes -0.0045% lower. The found total difference between the highest and lowest scores is equal to -0.446895% stock return per month.

The results of the portfolio analysis with the Five-Factor model control variables, found that the first hypothesis, where it was stated that a higher ESG score leads to a significantly different excess stock return, was accepted. It was found that on average, the stocks with the highest ESG score, tend to outperform the stocks with the lowest ESG score, by a 0.2755% higher stock return per month in the case of equal-weighted stocks. No statistically significant results for the value-weighted results were found. The second hypothesis, stating that a higher Environment score leads to a significantly different excess stock return, was rejected at the $p<0.05$ significance level for the equal-, and value-weighted returns. The third hypothesis, stating that a higher Social score leads to a significantly different stock return, was accepted for the equal-weight results. It was found that on average, the stocks with the highest Social score tend to outperform the stocks with the lowest Social score by a 0.25570% higher stock return per month. It was found that there are no statistically significant results in the case of value-weighted stocks. Finally, the fourth hypothesis stated that a higher Governance score leads to a significantly different stock return, which was rejected at the $p<0.05$ significance level for both the equal-, and value-weighted returns.

Keywords: ESG, Fama-French, Russel 1000

Table of content

Contents

Introduction

ESG investing and ESG scores are a concept from around 2005, that is nowadays gaining more and more traction within the investing community. In its early days, not much attention was devoted to the meaning behind it, and very few companies released information regarding their ESG scores. Between 2007-2010, many companies changed their views and started giving more and more information to the shareholders, to start tracking these scores. Eventually, these scores became more and more prominent to disclose and after a while enough data was gathered to start doing some research about it. ESG scores consist of Environmental, Social, and Governance scores, which are again based on various sub-factors. These three factors combined make the ESG score. ESG scores have given investors a measure to find firms that have a better and more sustainable business plan compared to their competitors in the same sector/industry. This score is generally used by investors to be more aware of what type of company they invest in. As stated by Admin (2020), it is found that if a company has a higher ESG score than its peers, it has scored higher in various individual factors of the ESG score. This can be in the Environment, Social or Governance score. Examples of sub-factors are having a higher number of women on the board, lower pollution emissions or having a higher employee diversity.

When you look at BlackRock, the world's biggest asset manager with an estimated 10 trillion Assets Under Management (AUM), an example of the increased importance of ESG scores is found. BlackRock has recently started to use its voting power (on behalf of its investors) to make various boards of directors more in line with the ESG standards. A specific example that is shown by Ambrose (2021) is the Exxon Mobile board vote, where BlackRock used its 6.7% stake together with various other “green” investors to vote for 2 new and more progressive, green directors. These recent changes show the ever-increasing importance of ESG scores for asset managers. Furthermore, knowing that BlackRock manages 10 trillion in assets, it can be expected that more moves from BlackRock will come to improve the ESG scores of companies that are not investing enough into their ESG scores.

When you look for control variables in regard to asset pricing, it is found that the strongest and most used control variables are the Fama & French (1992) Three-Factor Model. This model looks at five potential risk factors (market ß, size, B/M, financial leverage, and E/P^2) and found that size and value are the best proxies for explaining stock returns. Since 1992, many people have tried to expand this model even further. After more than 20 years, the same group of researchers came up with the newest asset pricing model, the Fama & French (2015) Five-Factor Model. Here, not only the Market ß, size, and B/M were used, but also the operating profitability and asset growth were included. This research paper will look into both models and try to find a new factor to expand the Fama & French Model. This paper will look at the possibility of adding the ESG scores as the fourth explanatory

factor. Furthermore, it will also look at the potential explanatory power of the Environment, Social, and Governance scores individually, to expand the Fama & French (1992, 2015) Models. While the same methodology as both previously named papers will be used in this research, this paper will specifically focus on the timeframe of 2010 till 2020. This is due to the limited availability of historical ESG data. As mentioned before, ESG investing and these scores were only starting to be tracked by some companies since approximately 2005, and it is further found that only since approximately 2010 most stocks actually received an ESG score in most databases. Amel-Zadeh & Serafeim (2018) found that investors are currently still using ESG scores motivated by financial reasons rather than ethical reasons. The question now becomes, if their motivation to follow ESG scores is correct. This leads us to the following main research question:

"Is there a difference between the stock returns of companies with a high ESG score compared to companies with a low ESG score?"

To conclude the introduction, a short explanation of all sections will be done. The paper starts with a review of the currently existing literature and the most important findings from other researchers. Furthermore, it explains why certain control variables are used and how this research paper is related to the existing literature. Afterward, several research hypotheses will be formulated that lay the foundation of this research paper. Next, the paper will have a section dedicated to the dataset. It will be discussed what timeframe is used, which index constituents are in the dataset, and how the data is gathered. Furthermore, the methodology section will consist of a further explanation of all the variables that are used. In this section, it will be explained how all values are calculated and what transformations have been made to the dataset. Then, the results section will consist of the findings of this paper, combined with the discussion, where the coefficients are interpreted. Furthermore, the potential shortcomings and weaknesses of the earlier found results will be examined critically. This paper will finish with the conclusion of the research.

Literature Review

The CAPM model

The Fama & French (1992) Three-Factor Model that is used in this research paper expands on the earlier found Capital Asset Pricing Model (CAPM). This is done by adding two additional risk factors: the "size"(SMB) and "value"(HML) effect. To understand this model, however, it is important to take a look at the CAPM model of Sharpe (1964) and Litner (1965) first, in order to understand the Fama & French (1992) model. The model of Sharpe (1964) and Lintner (1965) was the first to try to explain asset pricing. This model laid the foundation of the later found Three-Factor and Five-Factor Model of Fama & French (1992, 2015). A relationship between the systematic risk and expected stock returns was found and captured by the "Beta" of a stock. When the findings of the CAPM is turned into a formulated model, it becomes:

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i\,(E(R_{m,t}) - R_{f,t}) + \epsilon_{i,t} \quad (1)$$

First of all, it is important to explain the definition of all the variables used in the equation above. The first variable being used is $E(R_{i,t})$. This represents the expected return on investment on stock i, at moment t. Furthermore, the $R_{f,t}$ is the risk-free rate at moment t, usually taken as the theoretical rate of return on an investment with zero risk. Currently, there is no default investment that is totally "risk-free", it is more of a theoretical concept. However, the asset that is currently most often used as "risk-free" is the 1-Month U.S. Treasury Bill. For this research as well, the 1-Month U.S. Treasury Bill will be used to gather our "risk-free" rate. Next, the variable βi measures the volatility of stock i, in regards to the market risk premium. A positive relationship is found between stock returns and market risk. The paper of Sharpe (1964) suggests that a higher non-diversifiable market risk leads to a higher expected stock return. This suggests that investors are compensated for taking additional risk, for example, when they buy more volatile stocks. Finally, the definition of $E(Rm)$. This is the expected return on the market portfolio. It represents the average return of someone who invests in the total market.

Predecessors of Fama & French (1992) paper

After looking at perhaps the most important predecessor of the Fama & French (1992) Three-Factor Model, it is important to look at another paper that was written before the Fama & French (1992) paper. This paper was also looking to find a way to expand the CAPM model of Sharpe (1964) and Litner (1965), even before Fama & French succeeded at it. The research paper of Banz (1981) looked at the relationship between risk-adjusted returns and the total market value of NYSE common stocks

between 1926 and 1975. It was found that smaller firms have had higher risk-adjusted returns on average, when they were compared to larger firms. This was called a "size effect" in the paper of Banz (1981). It is easy to find the reason why this is seen as a predecessor to the "Small Minus Big" (SMB) factor of the Fama & French Three-Factor Model (1992). A similarity between the two papers is that the found "size" effect was predominantly found for small firms, while for average-sized and large-sized firms little difference in returns was found. There are, however, also important differences between these papers. In the paper of Banz (1981), it was stated that the size effect is not linear in market value. When this is compared to the paper of Fama & French (1992), it can be seen that there are no such statements being made. Another difference between the papers is also found in the methodology of both papers. The paper of Fama & French dissected these "size" effects by creating "portfolios", where you would look at the difference between the stock returns of the bottom quintile of firm sizes, compared to the stock returns of the top quintile of firm sizes. This was a brand new approach compared Banz (1981), which is why it revolutionized the asset pricing models.

The third and final research paper that is a predecessor of Fama & French (1992) is that of Chan et al. (1991). This paper is comparable to that of Fama & French (1992) in the way it is set up. Both papers start with a similar approach: both papers started with the creation of portfolios, consisting of various variables that might have some explanatory power for the excess stock returns of assets. When looking more specifically at the exact variables being used, the differences start to be noticed. The paper of Chan et al. (1991) looked at four variables: earnings yield, size, book-to-market ratio, and cash flow yield. Here, it can be concluded that the paper of Fama & French (1992) did have some similarities, namely the "size" and "book-to-market" factors. There are also a big differences between the papers, since there is no mention of an earnings yield or cash flow yield in the research of the latter. Another difference between both papers is that the paper by Chan et al. (1991) looks at the cross-sectional differences in returns on the Japanese Stock market, while the paper of Fama & French (1992) focuses on the U.S. stock market. Furthermore, the research of Chan et al. (1991) looks at a much shorter time period, from 1971 until 1988, compared to the research of Fama & French, which looks at a time period of 1962 until 1989. Finally, it is now known that the research of Chan et al. (1991) was on the right path: the book-to-market ratio and size factors were also found to be statistically significant in the paper by Fama & French in 1992. Differences were, however, that Chan et al. (1992) found that the book-to-market ratio and cash flow yield had the most significant positive impact on expected returns.

Fama & French (1992)

Now that some related academic findings are discussed, it is time to take a more in-depth look at the papers on which the control variables are based and how this paper is looking to expand the Fama & French (1992, 2015) Models. Starting with the first Fama & French (1992) paper, this counts as one of

the foundations of the existing asset pricing literature. The paper looked at various factors to expand on the CAPM model. After looking at their initial five potential risk factors (market ß, size, B/M, financial leverage, and E/P^2), they found that the 'size' and 'value' factors are the best proxies for explaining stock returns. When converting these findings of the Three-Factor Model, the following formulated model is found:

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i (E(R_{m,t}) - R_{f,t}) + s_i SMB_{t-1} + h_i HML_{t-1} + \epsilon_{i,t} \quad (2)$$

Here, there are some of the same factors as in the CAPM model. Yet again, the $E(Ri)$, rf, $E(Rm)$, and βi are there. These still have the same function as in the CAPM model. However, the Three-Factor model has this name because it expands the CAPM (one factor) model with two new factors: Small Minus Big (SMB) and High Minus Low (HML).

The factor SMB (Small Minus Big) is the 'size' factor, where the finding that small firms have a higher excess return is used. This is done by separating the companies into multiple quintiles, where you can then look at the difference in stock returns between the smallest and biggest quintile. This was done by Fama and French (1992) by retrieving the beta, SMB, and HML factors by estimating the regression equation, creating 25 portfolios with differences in the size and book-to-market values. More information on how to exactly create the portfolios using the Fama & French (1992) methodology can be read in the methodology section of this research paper.

The factor HML (High Minus Low) is the other new factor used by Fama and French (1992). This has a similar setup, where the finding is used that firms with a high book-to-market ratio have a higher excess return. Higher book-to-market ratio firms are known as 'value' stocks. Similarly, lower book-to-market ratio firms are known as 'growth' stocks. Furthermore, Fama & French (1992) decided to leave out the financial companies from the dataset. This was due to the relatively high leverage existing within these companies. Leaving those companies in the dataset would lead to a bias.

This paper will use these Three-Factors as control variables for the ESG scores. This is because according to Rehnby (2014) the Fama & French Three-Factor Model has a higher explanatory power than the CAPM model. Furthermore, he also states the Fama & French Three-Factor model is the most frequently used and widely known applied asset pricing model. This research paper will look at the options to expand this Three-Factor Model with the ESG scores as the fourth explanatory factor. Also, it will address the Environment-, Social-, and Governance scores individually to test the validity of these factors to expand the Three-Factor Model. In order to find more robust results for our hypotheses, this paper will not only be making portfolios, but it will also be using the methodology of the paper of Fama and MacBeth (1973) to determine the feasibility of expanding the Fama & French (1992) Three-Factor model with one of the ESG-, Environment-, Social- and Governance score factors. In the 1973 paper, the Fama–MacBeth regression is first introduced. This is a two-step procedure used to estimate the parameters of asset pricing models, such as the capital asset pricing

model (CAPM). The method consists of estimating the betas and risk premia for any factors that are expected to determine asset prices.

The two general steps in the Fama-MacBeth regression are as follows: first, one regresses each of the n asset returns against the m proposed risk factors to determine each asset's beta exposures. The second step is to regress all asset returns for each of the T time periods against the previously estimated betas to determine the risk premium for each factor. After the Fama-MacBeth regression, coefficients and p-values are found that will show how much, if any of the ESG-, Environment-, Social- or Governance scores are good predictors of excess stock returns.

Fama & French (2015)

Not only will this paper look at the Fama & French Three-Factor (1992) model to expand, it will also take a look at their newer research paper, the Fama & French Five-Factor (2015) model. This will be done in the second part of the research paper, when this paper will create the portfolios, to further inspect the possibilities to expand the Fama & French (2015) Five-Factor Model with the ESG-, Environment-, Social- or Governance score factor. The Five-Factor model adds the factors of Robust Minus Weak profitability (RMW) and Conservative Minus Aggressive investing (CMA) to the Three-Factor model. In this research a similar methodology to the Fama & French (2015) paper will be followed, by isolating the top and bottom quintiles of various variables. This research, just like the Fama & French (1992, 2015) papers, focused on the U.S. stock market. In the Fama & French papers, the data was gathered from CRSP. When converting the findings of the Five-Factor Model, the following formulated model is created:

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i\,(E(R_{m,t}) - R_{f,t}) + s_i SMB_{t-1} + h_i HML_{t-1} + r_i RMB_{t-1} + c_i CMA_{t-1} + \epsilon_{i,t} \quad (3)$$

Papers about the effect of ESG scores on stock returns

The reason for looking at ESG scores to improve the Three-Factor Model and Five-Factor Model is that ESG scores are a new and upcoming investing concept, with growing importance. Furthermore, it is only since a few years that there are enough data points to gather a meaningful interpretation and significance of the coefficients. As stated before, ESG scores have only been widely available since approximately 2010, where since 2005 more and more companies started to disclose information for analysts to calculate these scores. With this research, that goes as far as almost the most recent datapoints available, an attempt will be made to expand and update the existing literature using the newest datapoints. The relatively limited research that has been done so far on the ESG scores and their effect on excess stock returns has yielded the following results.

First of all, La Torre et al. (2020) found that highly ESG-rated firms report higher excess returns and lower volatility. La Torre et al. (2020) gives us a great indication of what results can be

expected from this research. Nevertheless, their research is very different in setup. First of all, La Torre et al. (2020) used very different control variables compared to this research. This paper will try to use more conventional control variables by looking at the Fama & French (1992, 2015) control variables. Furthermore, La Torre et al. (2020) used a relatively small dataset of only 46 companies in the Stoxx 50. In this research the Russel 1000 index will be used to look at all the companies that were in the index between 2010 and 2020 to get a more complete and representative dataset. This is a much bigger sample, which together with the different control variables leads us to expect stronger results.

Another paper that looked at ESG scores is that of Velte (2017). This paper looked more broadly at the effect of a higher ESG performance on the company's financial performance. Here, it was found that the Economic Social and Governance Performance (ESGP) has a positive impact on the Return On Assets (ROA). This effect on the ROA was found by looking at a sample selection of companies listed on the German Prime Standard (DAX30, TecDAX, MDAX) for the business years 2010-2014. This research will look at a longer time frame with more companies in it. In this paper, there are ~1500 companies included. Another difference is in the used control variables. The paper of Velte (2017) uses R&D expenses, firm size, firm unsystematic and systematic risk, and industry variables as controls, while this research will use the Fama & French Three-Factor Model and Five-Factor Model.

When analyzing the findings of the paper of Nollet et al. (2016), it can be seen that this paper looked at the effect of Corporate Social Responsibility and found that in the long term a U-shaped relationship exists, but only between the governance sub-component and Corporate Financial Performance (CFP). This was however found in the early days of sustainable investing, between 2007-2011, and our research will focus on the later period of 2010 until 2020. This paper will not only look at the total ESG scores, but also dive in deeper and look at the individual E, S, and G scores. The paper of Nollet et al. (2016) implicates that the strongest results are found for the governance sub-component. However, it is important to keep in mind that the coefficients and significance of these factors might have changed between the timeframe of Nollet et al. (2016) and the timeframe of this research paper, perhaps leading to different results in our dataset. Furthermore, the difference in used control variables could also lead to different results.

Finally, looking at the findings of Cornell (2020), it is stated that a high ESG rating can function as a hedge against climate shocks and unexpected changes in environmental regulation. This is an advantage, but it also comes at a cost according to the researcher. It is stated that these benefits come with a cost in the form of lower expected returns for investors. This is the other side of the ESG scores, where it could potentially cost investors money to invest in them. It still needs to be researched if this is the case for the Russell 1000 index. These previous research papers lay the foundation of our first and main research question.

Main research question and main hypotheses

RQ_1: How does a higher ESG score affect the excess stock returns of a company in the Russell 1000 during the period of 2010 to 2020?

This main research question will be tested using the following hypotheses:

H_0: Higher ESG scores have no effect on stock returns

H_a: Higher ESG scores have a significant effect on stock returns.

This hypothesis will be tested using the Fama-MacBeth regression on the Three-Factor model, and using portfolio analysis on the Fama & French Five-Factor model (2015). This paper will look at the alpha to find out if it explains significantly more of the variation in the stock returns. The other three hypotheses are stated in the section 'Research sub-questions & explanation' and will be further explained using the relevant literature below.

Research sub-questions & explanation

For the second, third, and fourth hypotheses, the ESG scores will be dissected into their individual pillars. Besides looking at the effects of ESG scores on stock returns in this paper, attention will be paid to the three other hypotheses. The effects of the Environment-, Social-, and Governance Scores on stock returns will be analyzed individually. This section starts by looking at the existing literature about the Environmental score on stock returns.

Several researchers have already touched upon the effect of Environmental scores on stock returns. Existing literature that is well peer-reviewed is relatively scarce. Most data is only available since approximately 2010, which is why this paper is one of the relatively few papers on this subject at this moment.

Looking at the paper of Velte (2017), specifically at his part of the effects found between the Environmental score and financial performance, it is found that these results are much weaker compared to the ESG scores. Furthermore, Velte (2017) states that the effect of the Environmental score is still positive and significant. However, it is also stated that governance performance has the strongest impact on financial performance (proxied by Return On Assets) in comparison to environmental and social aspects.

Another paper that looks at the Environmental score and financial performance is the paper by Derwall et al. (2005). This is an older paper that focused on the 'eco-efficiency premium' puzzle. Here, research was done to find the impact of the environmental aspect of CSR (Corporate Social Responsibility, a predecessor of ESG investing) on Tobin's q ratio and a firm's return on assets

(ROA). They found a positive and asymmetric relationship between eco-efficiency and the financial performance of a company. They found that firms with a high eco-efficiency score do not have a higher ROA than the control group. However, the more important finding of this paper is that firms with a low eco-efficiency score had a considerably lower ROA. This could imply that a higher Environment score leads to significantly higher excess stock returns.

Next, a look will be taken at the findings of Freedman & Patten (2004) on the effect of Environmental scores on stock returns. Here it is found that companies with a worse pollution performance (meaning a higher level of size-adjusted toxic releases) on their own, suffered more negative market reactions. However, it was found that when companies had more extensive environmental disclosure in their 10-K report, these disclosures mitigated the impact of their bad pollution performance. This paper will focus specifically on the sub-question of the effect of the individual Environmental score on excess stock returns.

Finally, looking at the findings of Albertini (2013), it is stated that there is more nuance to the positive effects. Taking it even further, there could be negative effects to having higher Environment scores. In Albertini's (2013) paper, it is stated that there could be positive effects to make investments in environmentally friendly production, but that the companies also need to make significant investments to achieve these goals without the costs being attributed directly to it. Furthermore, its stated that it is not always possible to pass all these costs back to the consumers, hurting the profitability of these firms. All these papers' previous findings lead to the second research question:

RQ_2: How does a higher Environment score affect the excess stock returns of a company in the Russell 1000 during the period of 2010 to 2020?

This sub-question will be tested using the following hypotheses:

H_0: Higher Environmental scores have no effect on stock returns

H_a: Higher Environmental scores have a significant positive effect on stock returns

Next, the third hypothesis of this research paper will find out if the addition of the Social score to the Fama French Three-Factor and Five-Factor models leads to a statistically significant coefficient. For the third hypothesis, this paper will take a more in-depth look at the panel, specifically looking for the Social score. Again, there is not much peer-reviewed research conducted on this factor in combination with ESG investing.

The findings of Velte (2017) again lay the foundation of our hypothesis. This paper states that the effects between the Social score and financial performance are much weaker compared to the ESG scores in general, but still statistically significant. Since the paper of Velte (2017) has already previously been discussed in this paper, the paper will continue with the findings of Mănescu (2011).

This paper looked at seven environmental, social and governance (ESG) individual attributes, to find out their influence on risk-adjusted stock returns. It also specifically looked at Socially Responsible Investing (SRI) and focused on the concept of "eco-efficiency" which can be thought of as the economic value a company creates relative to the waste it generates. It was found that Socially Responsible Investing produced statistically significant superior performance. This could be an indication for the findings of this paper, looking at the individual effect of a higher social score on stock returns. This paper will try to add to this research by looking at the Social score individually, instead of using SRI as a measure.

Finally, this paper will look at the paper of Brammer et al. (2006). This paper writes about Corporate Social performance and stock returns. This paper tries to reason for both sides of the argument, meaning the paper lets us think critically about what side is correct. The paper gives no final decision on whether or not Social investing leads to an increase or decrease in stock returns. One important argument for higher stock returns stated by Brammer et al. (2006), is that the firms must have a social conscience since it helps to satisfy their stakeholders. On the other side, the main finding of the paper is that firms with higher social performance scores achieve lower returns. Furthermore, it is stated that firms with the lowest possible CSP scores of zero outperformed the market. All these papers' previous findings lead us to the third research question, which is as follows:

RQ_3: How does a higher Social score affect the excess stock returns of a company in the Russell 1000 during the period of 2010 to 2020?

This sub-research question will be tested using the following hypotheses:

H_0: Higher Social scores have no effect on stock returns

H_a: Higher Social scores have a significant effect on stock returns

Finally, the fourth hypothesis will attempt to find out if the addition of the Governance score to the Fama & French Three-Factor (1992) and Fama & French Five-Factor (2015) model is statistically significant. There are yet again only few papers that looked at the ESG scores and their effects on stock returns.

When looking at the existing knowledge, the paper of Velte (2017) again found some important results. This paper looked specifically at the effects found between the Governance score and financial performance. This paper found the strongest effects on financial performance for the Governance factor individually and it is possible that similar findings will be gathered in this paper. A difference was that Velte (2017) looked at the German stock exchange specifically. In Germany, the Governance score is focused on relatively much compared to the U.S. stock market, due to heavier regulations existing in Germany compared to the Environmental and Social scores.

Furthermore, looking at the findings of Mănescu (2011), it is stated that when looking at a few of the specific factors that make up the Governance score and their effect on stock returns, no significant relationship is found. Our paper will go one step back and look at the Governance score altogether and try to find if there is a significant relationship between the total Governance score and stock returns. The findings of Mănescu (2011) make it less likely that the Governance score in total would have a positive effect on excess stock returns.

Finally, a look will be taken at the paper of Bozec & Bozec (2011). This paper looks at many different studies that were conducted on the Governance-Performance relationship in the existing literature. The main findings were that many of the studies based on U.S. and Canadian markets show conflicting results. Half of the U.S. and Canadian studies that were looked at by Bozec & Bozec (2011) suggested that there is no relationship or a negative relationship between governance and performance, while the other half generally supports a significantly positive link between the two. The paper explains that the absence of a clear relationship between the overall Governance and performance in North America might be interpreted as a sign that Canadian and US firms have already adopted relatively good governance practices. This makes that any additional improvements in overall governance have no significant impact on their performance. This means that it is still relatively unclear what effects could be found in this paper. The various findings in the research papers discussed so far lead to the fourth research question being stated as follows:

RQ_4: How does a higher Governance score affect the excess stock returns of a company in the Russell 1000 during the period of 2010 to 2020?

This sub-question will be tested using the following hypotheses:

H_0: Higher Governance scores have no effect on stock returns.

H_a: Higher Governance scores have a significant effect on stock returns.

Data

For this research paper the focus is on the period of 2010 until 2020. Furthermore, this paper looks to focus on the companies in the Russell 1000 index specifically. This is because this provides us a bigger dataset than the S&P 500 and this dataset contains more 'small' companies. This index was chosen because the 'size' effect discussed in the literature review is found to be the strongest in the smaller companies. Furthermore, stronger results can be expected by using a bigger dataset. The Russell 1000 is a subset of the larger and perhaps more well-known Russell 3000 index. Looking specifically at the differences in total U.S. market coverage, it is found that the Russel 1000 index comprises approximately 92% of the total U.S. market capitalization, according to Ganti (2022). If you compare this to the approximately 75% coverage by the S&P 500, according to Thune (2022), it is much more.

Furthermore, this paper will gather its data from the Refinitiv EIKON dataset. This is a data provider that is mainly used to analyse financial information, and the paper will follow the definitions of this program in our research. The other data that is used comes from the Kenneth R. French Data Library. In this data library the daily, weekly, monthly, and yearly Fama & French factors are tracked for various regions of the world and uploaded onto different Excel sheets. These factors are used as control variables in our Five-Factor model. The factors given in Kenneth R. French Data Library are for the general U.S. Market, the NASDAQ, NYSE, and, AMEX, instead of being made specifically for the Russell 1000. As mentioned before, Russell comprises approximately 92% of the total U.S. Market cap, so the expectation is that these differences will not be very influential.

After the selection of companies, the next step is to actually gather the data through Refinitiv EIKON. The following variables are gathered from EIKON (between brackets '()' is their EIKON name): The ESG score ('TRESGS'), Environment score ('ENSCORE'), Social score ('SOSCORE'), Governance score ('CGSCORE'), the book value outstanding shares fiscal ('WC05491'), the common shares outstanding ('WC05301'), the market capitalization ('MV'), the historical beta ('897E'), a date variable (automatically implemented), the ICB Industry Name ('ICBIN') and finally the stock returns (combination of 'P#T' & 'RI'), where the stock returns take into account the paid out dividends and issuance of new shares. Some of these factors still need some editing before they can be used, which is explained in the methodology section.

Descriptive statistics, matrix of correlations & ESG score construction

Looking at the descriptive statistics of the gathered data, it is possible to check if the data looks as expected. First, it is noticed that all variables have at least 121.173 observations in them. Furthermore, the average Governance and Social score is much higher for companies, compared to their

Environmental score. Next, it is found that the mean of the monthly return is 0.5%, which seems reasonable as an average monthly return, since this implies a 6% yearly stock return. Furthermore, the minimum and maximum of the beta are very wide-spread. This is due to new companies being added to the Russell 1000 experiencing a lot of additional 'first month' volatility. To not have this first month influence the research too much, the decision was made to exclude outliers with a beta >50 and <-50. Before this, beta's of -2700 and +1300 were present in the data.

Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
ESG score	121273	46.015	19.968	.47	95.15
Environment score	121197	35.755	28.594	0	98.55
Social score	121173	48.594	21.46	.26	97.95
Governance score	121225	51.371	22.857	.25	99.56
Beta	152618	1.157	.917	-45.361	42.5
LN Stock return	136374	.005	.123	-3.401	4.868
LN market capitalisation	159544	8.666	1.498	-4.605	14.607
LN book-to-market	122756	-.941	.918	-7.343	3.728

Figure 1: This table summarizes the number of observations (Obs) per variable. Next, it shows the mean per variable, over all observations. Furthermore, the standard deviation of this mean is given. In the last 2 columns, the minimum and maximum values for all used variables are stated.

After the descriptive statistics, it is also possible to look at the correlations between our main variables of interest. The most important finding is that ESG-, Environment-, Social- and Governance scores are relatively highly correlated with the natural logarithm of the market capitalization. Looking at other research papers that had similar findings, Akgun et al. (2021) found a similar pattern. In this paper, the correlation is mainly explained by the fact that ESG scores tend to have a size bias toward larger companies. The most significant effect was found in the large cap stocks. Furthermore, they stated that average ESG scores generally rise as you climb the market cap spectrum, with the largest companies scoring higher than their smaller peers within the large cap universe. Furthermore, our own addition to these findings is that larger companies are much more likely to get an ESG-, Environment, Social-, and Governance score. This is due to the availability of the amount of information that is needed to make up the ESG scores. For smaller companies, there is not always enough information disclosed to calculate these ESG scores, leading to a size bias for ESG-, Environment, Social-, and Governance scores.

Matrix of correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) ESG score	1.000							
(2) Environment	0.874	1.000						

score								
(3) Social score	0.880	0.730	1.000					
(4) Governance score	0.691	0.435	0.384	1.000				
(5) Beta	-0.114	-0.103	-0.094	-0.082	1.000			
(6) LN stock return	0.000	-0.005	0.006	-0.007	0.010	1.000		
(7) LN market cap.	0.524	0.512	0.494	0.270	-0.286	-0.001	1.000	
(8) LN book-to-market	-0.099	-0.049	-0.139	-0.015	0.138	-0.011	-0.266	1.000

Figure 2: This table summarizes the correlations between the main variables that are used in this research paper. In the first row, the overall correlations between the ESG score and the other variables used.

In order to understand what ESG scores are and how they are made up, it is important to look at how they are created. This is done by independent agencies, based on the publicly available (annual reports, corporate websites, CSR papers, and news sources) information. These agencies then decide on the importance of these events and are then also scored by the same independent agencies. The paper of Hulsebosch (2020) shows the Environment-, Social-, and Governance scores and their weight.

ESG score construction				
Pillar	**Category**	**Indicators in rating**	**Weights**	**Pillar weights**
Environment	Resource use	19	11%	
	Emissions	22	12%	34%
	Innovation	20	11%	
Social	Workforce	29	16%	
	Human Rights	8	4.50%	35.50%
	Community	14	8%	
	Product Responsibility	12	7%	
Governance	Management	34	19%	
	Shareholders	12	7%	20.50%
	CSR Strategy	8	4.50%	
Total		178	100%	100%

(adapted source: Hulsebosch (2020))

Figure 3: This overview shows the construction of the ESG scores in 2020. There is not much information available about the composition of the ESG scores, thus this scoresheet from 2020 is used. The weighting columns show how much every pillar and sub-section of a pillar influence the total ESG score.

Methodology

In this research paper, the methodology of the Fama & French (1992) and the Fama & French (2015) will be followed. In these papers, the setup to find the Three-Factor and Five-Factor Model is very similar. The Fama & French approach expands on the CAPM model, by deriving factors that tend to explain the variations in cross-sectional stock returns. This paper will create a multi-factor financial model, to explain market events and or equilibrium asset valuations. This research will use Sharpe's (1964) CAPM methodology, further expand that with the Three-Factor and Five-Factor Model following the Fama & French (1992, 2015) methodology, and finally look to expand this model with our own new factor, the ESG-, Environment-, Social- or Governance score, to see if it has any statistically significant explanatory power.

The first step to take in our research, is to gather the ISINs from all the companies that were in the Russel 1000 Index, during the time period of 2010 to 2020. Any company that was part of the index during one or more of those years is included in the list. From these yearly lists the duplicate ISINs will be removed, together with any duplicates due to the existence of multiple share classes. Only one of those classes of shares stays in the dataset. This is necessary because class ‘A’ and ‘B’ shares are very similar in terms of characteristics. Keeping both shares in the dataset will cause companies with multiple share classes to be overrepresented.

For the second step, all the companies that are in the ‘financial’ industry will be filtered out. This has been done in various other research papers, like Fama & French (1992, 2015), but is explained further in Dirkx & Peter (2020). Here, it is stated that financial companies are left out because their book value is much more leveraged than the other industries. This would influence the coefficients of the book-to-market control variable. This can be done by using the ICB Industry Name variable of the Refinitiv EIKON database. This variable classifies every stock under one of the 10 industry types, based on which one it relates most to. In our research, this caused our dataset to go from 1593 companies, back to 1398 companies. This resulted in a filtered dataset, but without an upwards bias in the book-to-market ratio.

Finally, with these edits, the gathered the list of companies could be used in the research. First of all, the stock returns variable will be used to track the returns of the individual companies. Furthermore, the control variable SMB will be used. SMB stands for Small Minus Big, meaning that you control the stock returns for the on average higher returns that small companies make. Next is HML, this is also a control variable. HML stands for High Minus Low, meaning that you control the stock returns for the on average greater returns that high book-to-market companies make. These high book-to-market companies are known as ‘value’ stocks. Furthermore, the stock returns are controlled for CMA. CMA stands for Conservative Minus Aggressive investment. It has been found in previous

research, like the Fama & French (2015), that Conservatively investing companies tend to outperform Aggressively investing companies. Finally, the stock returns are controlled for RMW. RMW stands for Robust Minus Weak. Like with the other control variables, it is found that the research of Fama & French (2015) has found that companies with a Robust operating profitability outperform the companies with Weak operating profitability.

The first change made to the dataset is to make all variables lagged relative to the stock returns. Looking at the ESG score at point t for example, you must look at the stock returns of the month following that variable, meaning the returns of month t+1. The only different lag is with book-to-market, this variable will be lagged by 6 months. This is done because the book-to-market value only changes once every year, so it is important to look back further than 1 month to have the same 'lagged' effect. Furthermore, some individual variables need some editing before they can be used.

The first variable that needs some editing is that of the stock returns. The stock returns variable used in this research, the Return Index ('RI'), also takes into account dividend payments and share issuance. Following the Fama & French methodology, it is important in our dataset to control our stock returns for the Risk-free rate. This is done based on the 1-month U.S. Treasury Bill. This value is subtracted from every month's stock return. Furthermore, in order to derive the monthly stock returns from the RI, the following calculation needs to be done:

Stock return at moment t $\approx LN(RI_{t+1}) - LN(RI_t)$ (4)

This stock return is used in the rest of our research. Furthermore, the P#T variable is used to see when a company is delisted or had a merger with another company. The Return Index (RI) keeps copying the last value if a company stops existing, while the P#T value keeps the cell empty. Copying the same value while a company is delisted would cause our return coefficient to be biased towards 0, so a combination of these variables should be made. Also, the natural logarithm of the stock returns is used because the natural logarithm assumes constant compounding of the returns and is slightly more accurate. It is also important to edit some of the control variables. First of all, the gathered 'book-value outstanding shares fiscal' needs to be turned into the book value:

Book-value = Book-value outstanding shares fiscal * Common shares outstanding (5).
Book-to-market value = Book-value / Market capitalization (6).

Again, you take the natural logarithm of the book-to-market ratio. This has several advantages, like having all control variables in the same format. By taking the natural logarithm, the negative book-to-market ratios are deleted, since it is impossible to take the natural logarithm of a negative number. By doing this, the same methodology as both Fama & French (1992, 2015) research papers is followed.

Furthermore, we also take the natural logarithm of the variable market capitalization. After these conversions, the dataset is finally ready to be imported into STATA, where the analysis can be started.

Now, it is possible to finally look at the various regressions that will be done. Two types of regressions will be done in order to find the results. The first option will first make use of the Fama MacBeth regression, first conducted by Fama & MacBeth (1973). The Fama MacBeth regression is a regression consisting of two parts. It allows us to control the set of variables used in our research when researching the main variable of interest, ESG scores. The first step of the Fama MacBeth regression is the periodic cross-sectional regression of the dependent variable of interest, denoted as Y, on one or more independent variables, X1…Xn. Here, to find the cross-sectional regression coefficients, it is important to look at the individual data from each time period t and run the regressions. This is done by the following general formula:

$$R_{i,t} = \delta 0 + \delta_{1,t} X1_{i,t} + \ldots + \delta_{n,t} XN_{n,t} + \varepsilon_{i,t} \quad (7)$$

Now, for the second step, we look at all the time periods together. This can be done by taking the average of each of these periodic regression coefficients δ. Then, one can conclude what the effects are, over the entirety of the time series. The combination of these two steps are how the Fama MacBeth regressions get their coefficients.

This research paper will also look at a second route to find the answers to our research questions. This is done by making our own portfolios and adding our ESG-, Environmental-, Social-, or Governance score to the portfolio. Including both methods in this research paper will make the findings of this research more robust. The portfolios are made based on the relative height of the values of all the variables. For all of the used variables, except for the stock returns and market capitalization, they will be split into one of three groups. For example, for the ESG score, every time period is looked at individually and every value is put into one of the three groups. The first group is that of the top 33.3% highest ESG scores with their stock returns is denoted by the number 1. The next group is the middle 33.3%, denoted by a 2 in the dataset. Finally, the bottom 33.3%, denoted by a 3 in the dataset. After creating these groups, the difference in average returns between the top 33.3% and bottom 33.3% groups can be compared. After doing this, the “ESG score” factor will be found by looking at the difference in stock returns of the three quintiles. Then, the control variables of the Five-factor Fama & French model from the Kenneth R. French - Data Library will be used to see if the results are statistically significant. After that, the results are compared to the known Fama & French Five-factor model, to see if they have significantly more explanatory power and thus a significantly different alpha.

Lastly, to find the right results, all coefficients will be controlled for the Newey & West (1987) standard errors. This is due to the return variable and control variables which might be persistent. Even though two variables are probably only correlated for one period, they may be

persistent in multiple time periods. Without the Newey & West (1987) standard errors, the hypotheses might be supported too easily. This paper will control the results for 6 periods of lags. At the end of the research, a conclusion will be drawn to potentially expand the Fama & French Three-Factor (1992) and Fama & French Five-Factor (2015) Models with one or multiple of the following factors. This will be done to expand the Fama & French Five-Factor Model below. The factor ESG stands for the total ESG score factor, ENV stands for the Environment score factor, SOC stands for the Social score factor, and GOV stands for the Governance score factor.

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i\,(E(R_{m,t-1}) - R_{f,t-1}) + s_i SMB_{t-1} + h_i HML_{t-6} + r_i RMB_{t-1} + c_i CMA_{t-1} + t_i ESG_{t-1} + \epsilon_{it} \quad (8)$$

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i\,(E(R_{m,t-1}) - R_{f,t-1}) + s_i SMB_{t-1} + h_i HML_{t-6} + r_i RMB_{t-1} + c_i CMA_{t-1} + e_i ENV_{t-1} + \epsilon_{it} \quad (9)$$

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i\,(E(R_{m,t-1}) - R_{f,t-1}) + s_i SMB_{t-1} + h_i HML_{t-6} + r_i RMB_{t-1} + c_i CMA_{t-1} + s_i SOC_{t-1} + \epsilon_{it} \quad (10)$$

$$E(R_{i,t}) - R_{f,t} = \alpha_i + \beta_i\,(E(R_{m,t-1}) - R_{f,t-1}) + s_i SMB_{t-1} + h_i HML_{t-6} + r_i RMB_{t-1} + c_i CMA_{t-1} + g_i GOV_{t-1} + \epsilon_{it} \quad (11)$$

Results and Discussion

Fama-MacBeth regression on hypothesis 1,2,3 & 4

As stated in the methodology, there are two types of regressions done in this research. The Fama & MacBeth (1973) regression is the first analysis that will be done to find the effects of ESG scores on excess stock returns. In this part, a Fama & MacBeth regression with the Three-Factor Model control variables is done. The control variables are the historical beta, the natural logarithm of the market capitalization, and the natural logarithm of the book-to-market ratio. When you look at figure 1, the first column shows the found coefficients of the Fama & MacBeth (1973) regression on the total ESG score, the second column on the Environmental score, third on the Social score, and the fourth on the Governance score.

Looking at the results of the Fama & MacBeth (1973) regression in the first column, it can be found that the coefficient of the ESG score on the monthly return is very small and also statistically insignificant. The coefficient of -0.000027, therefore, does not have any meaningful interpretation and cannot be used to predict the cross-section of stock results. Also, it is found that none of the control variables are statistically significant either in the first column. This means that the control variables cannot be interpreted either. The results of the second, third, and fourth columns respectively, state the coefficients of -0.000041, 0.000012 and -0.000045. Since the hypotheses are only accepted when it is significant at the 5% level, the second and third hypotheses are rejected. This insignificance is also the case for all the control variables in column 2, 3 and 4. It would be reasonable to expect to find more significant Beta's, book-to-market ratios, and market capitalizations for our control variables since these variables are already well established. This means that these found coefficients do not have any meaningful interpretation, contrary to the generally accepted asset pricing literature. Finally, for our fourth hypothesis, a statistically significant result is found. This coefficient can thus be interpreted. The coefficient of the Governance score means that on average, in our sample, when the Governance score is one point higher, the average monthly stock return becomes -0.0045% lower. Looking at descriptive statistics, this means that the total difference between the highest and lowest scores is equal to a -0.446895% stock return per month. It can be concluded here that a higher Governance score has a significantly negative effect on stock returns.

This result is in contrast with most of the previously discussed literature. Putting more focus on research papers with similar expectations, it can be seen that the paper of Bauer et al. (2004) also expected a higher Governance score to lead to higher stock returns. His findings, however, were also the opposite. In the discussion of his research paper, Bauer et al. (2004) stated that this finding might be due to current corporate governance standards having, to a large extent, already been incorporated into stock prices. This research paper, however, focused on the differences between the UK stock market and the European Monetary Union (EMU). Looking at the implications of this paper on the

U.S stock market, an argument could be made that a higher Governance score could actually lead to higher excess stock returns, but that this is already priced into the stock price too much, leading to the actually negative coefficient that were found for hypothesis four. Furthermore, it can be seen that the constant, also known as the alpha, is not statistically significant in any column.

Name	stock returns ESG	stock returns ENV	stock returns SOC	stock returns GOV
ESG score	-0.000027			
	(-0.864898)			
Environment score		-0.000041*		
		(-1.728351)		
Social score			0.000012	
			(0.569767)	
Governance score				-0.000045**
				(-2.011148)
Beta	-0.002002	-0.001989	-0.002065	-0.002015
	(-0.889297)	(-0.886105)	(-0.915251)	(-0.898310)
LN market cap.	0.000845	0.001108	0.000514	0.000851
	(0.972371)	(1.222946)	(0.630185)	(1.099941)
LN book-to-market	-0.001372*	-0.001283	-0.001431*	-0.001357*
	(-1.658318)	(-1.626790)	(-1.711236)	(-1.661686)
cons	0.000052	-0.001895	0.001127	0.001055
	(0.006085)	(-0.208611)	(0.131972)	(0.127587)
r2	.06251	.06323	.0614	.06237
r2_a	.05792	.05863	.0568	.05777
N	108972	108896	108872	108924
t statistics in parentheses				
="* p<0.10	** p<0.05	*** p<0.01"		

Figure 4: This first column showcases the found coefficients of the Fama & MacBeth (1973) regression on the total ESG score. The second column does this for the Environmental score, the third column for the Social score,

and the fourth column isolates the Governance score. On all columns, the variables are controlled for the historical beta, the natural logarithm of the market capitalization, and the natural logarithm of the book-to-market ratio. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r2_a), which is a corrected goodness-of-fit measure for linear models. It corrects the r2, because that tends to optimistically estimate the fit of the linear regression by always improving when a new variable is added. Finally, N represents the number of observations used in this research.

Portfolio analysis

The second type of regression done in this paper is the portfolio sort. Here, a long-short portfolio is created based on various characteristics. Our variables will be controlled using the Five-Factor Fama & French (2015) methodology, as was described in the methodology section. For every hypothesis, four tables will be shown. Two will be with the equal- and value-weight quintile results, and then two will show the equal- and value quintile difference results. For the first two tables of every hypothesis, it is important to keep in mind that the results are not controlled for autocorrelation. These shown results are purely informative to see how the variables behave and if the coefficients are negative or positive at every quintile. The most important variable to look at is the 'cons', also known as alpha. This variable shows the part of the stock returns that are not explained by the Fama & French Five-Factor Model (2015).

The equal- and value-weight ESG, Environment, Social, and Governance quintiles

For the first hypothesis, this paper will look at the three ESG-created Equal-Weight quintiles in figure 5a. Of these quintiles, it can be seen that the market return subtracted by the risk-free rate, the Small Minus Big, and the HML are all statistically significant control variables overall. This confirms that the Fama & French (1992) factors are also found to be significant in our research in combination with the ESG scores. Furthermore, it can be seen that only quintiles 1 and 2 of CMA have a significant effect on the ESG score quintiles. From the RMW factor, only the first quintile seems to have significant explanatory power in combination with ESG scores. Thus, we see more variations in our results and control variables when comparing this with the results of the Fama & French (2015) research paper. Still, we can conclude that most of the results in combination with our control variables are in line with the literature. Furthermore, looking at the 'cons', which is the alpha, it can be seen that the alpha does increase slightly upwards, the more you go towards quintile 3. This is similar to some of the literature that is analyzed before. The found coefficients will be interpreted in the next step of the research, looking at the differences between quintiles 3 and 1.

Looking at the three ESG-created Value-Weight quintiles, fairly similar results can be seen in figure 5b. The mkt-RF and SMB factors are statistically significant in all quintiles. For the HML

factor, only the third quintile is statistically significant. The value-weighted results of the Fama & French Three-Factor Model are a bit weaker, but still statistically significant coefficients in general. Looking at the CMA factor, all 3 quintiles are statistically significant control variables. Furthermore, focusing on the RMW factor, only the third coefficient is statistically significant. These findings are also fairly similar to the findings of the Fama & French Five-Factor Model (2015). The not statistically significant results are a surprising deviation from our expected results since these factors have been found in many other research papers. Some of the additional deviations compared to the equal-weighted results, could be due to the fact that value-weighted stocks are more biased towards big firms. Furthermore, looking at the 'cons', which is the alpha, it can be seen that the alpha does move upwards, the more you move towards quintile 3. This is similar to some of the literature that was analyzed before. This is also a similar finding to the equal-weight results that were just discussed. The found coefficients are interpreted in the next step, looking at the differences between quintiles 3 and 1.

Name:	EW return q. 1 ESG	EW return q. 2 ESG	EW return q. 3 ESG
Mkt-RF	1.119724***	1.112274***	1.031005***
	(31.782986)	(37.141598)	(42.960268)
SMB	0.576751***	0.462980***	0.186610***
	(8.923125)	(8.426673)	(4.238255)
HML	0.224277***	0.300301***	0.192351***
	(3.519411)	(5.543778)	(4.430988)
RMW	0.040098	0.141244*	0.199299***
	(0.434576)	(1.800846)	(3.170802)
CMA	-0.351188***	-0.255136***	-0.056121
	(-3.256867)	(-2.783531)	(-0.764020)
cons	-0.007882***	-0.006195***	-0.005127***
	(-5.575559)	(-5.155234)	(-5.324171)
r2	.9383	.9519	.9566
r2 a	.9359	.95	.9549
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 5a: This figure shows the average equal-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative

Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***

Name:	VW return q. 1 ESG	VW return q. 2 ESG	VW return q. 3 ESG
Mkt-RF	1.025937***	1.075998***	0.945990***
	(34.203182)	(46.936052)	(82.852936)
SMB	0.252677***	0.129512***	-0.145801***
	(4.591516)	(3.079289)	(-6.960269)
HML	-0.062395	0.009433	-0.080760***
	(-1.150001)	(0.227481)	(-3.910397)
RMW	-0.137299*	0.046384	0.203669***
	(-1.747717)	(0.772543)	(6.810897)
CMA	-0.388097***	-0.209208***	0.099119***
	(-4.227295)	(-2.981601)	(2.836318)
cons	-0.005124***	-0.005346***	-0.003751***
	(-4.256728)	(-5.811445)	(-8.186966)
r2	.9345	.9606	.9844
r2 a	.9319	.9591	.9838
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 5b: This figure shows the average value-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***

When the focus is shifted towards the three quintiles for the Environment equal-weight results in figure 6a, it can be seen that the market return subtracted by the Risk-free rate, the Small Minus Big, and the HML are all statistically significant control variables overall. This confirms that these Fama & French (1992) factors are also statistically significant in this research with the Environment scores. Furthermore, it can be noticed that only quintile 1 of CMA has a statistically significant effect on the Environment score quintiles. Also, for the RMW factor, only the second and third Environment quintile seems to have significant explanatory power in combination with Environment scores. In general, this means that the control variables of the Fama & French (2015) are not all statistically significant, but most of them are. All these factors add at least some explanatory value for some of the quintiles. Furthermore, looking at the 'cons', which is the alpha, it can be found that some unexplained downward movement happens between quintiles 1 and 2. Zooming out, it is still found that the alpha moves upwards towards quintile 3. The found equal-weight coefficients will be interpreted in the next step of the research, where the differences between quintiles 3 and 1 are analyzed.

Looking at the three Environment-created value-weighted quintiles in figure 6b, it can be seen that only the mkt-RF variable is statistically significant in all 3 quintiles. Looking beyond the mkt-RF factor, the SMB is statistically significant at both the first and third quintile, while HML is only statistically significant at the third quintile. This means that a bit more than half of the control variables of the Fama & French (1992) model are statistically significant. However, as stated before, all the variables have some explanatory power in one of the quintiles, so they are still important to have as control variables. Next, the results show that only quintiles 1 and 3 of RMW have a significant effect on the Environment score quintiles. Furthermore, it can be seen that this is a similar result with CMA, where it also only seems to have significant explanatory power in combination with ESG scores in the first and third Environment quintiles. In general, this means that the control variables are good and mostly in line with the literature discussed in the Literature Review section. Furthermore, looking at the 'cons', which is the alpha, almost no movement is seen between quintiles 1 and 3. This is in contrast with the literature reviewed before, where we saw a lot of research with positive or negative effects. In general, this shows that we could expect to find relatively weak results for the value-weighted portfolios. The found coefficients will be interpreted in the next step, where a look at the differences between quintiles 3 and 1 is taken.

Name:	EW return q. 1 env	EW return q. 2 env	EW return q. 3 env
Mkt-RF	1.105788***	1.124644***	1.032783***
	(30.720688)	(37.395447)	(42.053030)
SMB	0.551347***	0.499173***	0.175505***
	(8.348900)	(9.046894)	(3.895145)
HML	0.206059***	0.298372***	0.210834***
	(3.164832)	(5.484825)	(4.746014)

RMW	0.018579	0.185432**	0.176490***
	(0.197080)	(2.354212)	(2.743888)
CMA	-0.468224***	-0.107596	-0.088926
	(-4.250003)	(-1.168896)	(-1.183029)
cons	-0.006593***	-0.007219***	-0.005378***
	(-4.564739)	(-5.981602)	(-5.457432)
r2	.9338	.953	.9549
r2 a	.9311	.9512	.9531
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 6a: This figure shows the average equal-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***

Name:	VW return q. 1 Env	VW return q. 2 Env	VW return q. 3 Env
mkt-RF	1.010710***	1.066779***	0.946300***
	(34.065166)	(44.888778)	(80.849538)
SMB	0.262114***	0.050976	-0.126761***
	(4.815251)	(1.169163)	(-5.903082)
HML	-0.087167	-0.043355	-0.067671***
	(-1.624187)	(-1.008556)	(-3.196315)
RMW	-0.170236**	0.065423	0.186382***
	(-2.190754)	(1.051115)	(6.080095)
CMA	-0.449101***	-0.143254*	0.106815***
	(-4.945427)	(-1.969451)	(2.981658)
cons	-0.004110***	-0.004259***	-0.004193***
	(-3.452088)	(-4.466219)	(-8.926585)

r2	.935	.9546	.9838
r2 a	.9324	.9528	.9832
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 6b: This figure shows the average value-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***.

Taking a look at the three Social-created Equal-Weight quintiles in figure 7a, it can be seen that the market return subtracted by the Risk-free rate, the Small Minus Big, and the HML are all statistically significant control variables in all of the three quintiles. This confirms that the Fama & French (1992) factors are also found to be statistically significant for the Social scores. This is the result that was expected based on the literature. Looking at the Fama & French (2015) control variables, it can be noticed that only quintiles 1 and 2 of CMA have a statistically significant effect when we look at the Social score quintiles. Furthermore, the RMW factor only seems to have significant explanatory power in combination with Social scores in the third quintile. In general, this means that the Fama & French (2015) control variables are important in at least some of the quintiles. Thus, it is important to include them in our research and it makes our results more robust. Finally, looking at the 'cons', which is the alpha, it is important to notice that the alpha does increase when you look at the move from quintile 1 towards quintile 3. This is similar to some of the findings in the literature that has been analyzed before, and it follows some of the expectations from hypothesis 3. The found coefficients will be interpreted in the next step, where we look at the differences between quintiles 3 and 1.

When the focus is shifted to the three Social-created value-weighted quintiles in figure 7b, it can be seen the mkt-RF, and the SMB are both statistically significant control variables in the three quintiles. Furthermore, notice that the HML factor is statistically significant only in the third quintile. This confirms that these Fama & French (1992) factors are in general statistically significant in combination with the Social score quintiles. Looking at the new control factors of the Fama & French (2015) model, it can be seen that both CMA and RMW are statistically significant in all three quintiles. This means that these control variables are also found to be statistically significant in combination with the Social score factor. This further strengthens the research on the Fama & French

(2015) factors, showing their explanatory power in combination with the Social scores. Finally, looking at the 'cons', which is the alpha, it can be found that the alpha does increase when we look at the move from quintile 1 towards quintile 3. This is similar to some of the literature we have analyzed before, and it follows the expectations of hypothesis 3. All three 'cons' or alphas have high t-scores, meaning that they are statistically significantly different from 0. The found coefficients will be interpreted when we look at the differences between quintiles 3 and 1 in the next step of the research.

Name:	EW return q. 1 soc	EW return q. 2 soc	EW return q. 3 soc
Mkt-RF	1.123111***	1.103405***	1.037514***
	(30.922346)	(36.701811)	(43.963152)
SMB	0.573794***	0.459426***	0.193628***
	(8.610945)	(8.329398)	(4.472079)
HML	0.317368***	0.233891***	0.166511***
	(4.830747)	(4.300976)	(3.900663)
RMW	0.085494	0.144953*	0.153305**
	(0.898756)	(1.840931)	(2.480326)
CMA	-0.275296**	-0.307252***	-0.079749
	(-2.476427)	(-3.339055)	(-1.104067)
cons	-0.007539***	-0.006702***	-0.004982***
	(-5.172350)	(-5.555208)	(-5.260304)
r2	.9363	.9498	.9584
r2 a	.9338	.9478	.9567
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 7a: This figure shows the average equal-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***.

Name:	VW return q. 1 soc	VW return q. 2 soc	VW return q. 3 soc

mkt-RF	1.029744***	1.054568***	0.947363***
	(34.572109)	(49.282754)	(80.770619)
SMB	0.236280***	0.097615**	-0.140092***
	(4.323841)	(2.486455)	(-6.510199)
HML	0.057581	-0.005114	-0.093879***
	(1.068744)	(-0.132133)	(-4.424921)
RMW	-0.155135**	0.137051**	0.193479***
	(-1.988680)	(2.445462)	(6.298371)
CMA	-0.295629***	-0.221809***	0.097151***
	(-3.242802)	(-3.386688)	(2.706216)
cons	-0.005424***	-0.004853***	-0.003850***
	(-4.537973)	(-5.652149)	(-8.180019)
r2	.9365	.9632	.9836
r2 a	.934	.9618	.9829
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 7b: This figure shows the average value-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***.

Finally, focusing on the three Governance score quintiles in figure 8a, pretty similar results to that of the first, second, and third hypotheses can be seen. First of all, one can see that the Mkt-RF, SMB, and HML are statistically significant in all three quintiles. This shows that the Fama & French (1992) control variables are also statistically significant when looking at the Governance score quintiles. This further strengthens the Fama & French (1992) model, since it is shown that they are also significant in combination with the Governance score. Continuing the focus on the Fama & French (2015) factors, it can be seen that the RMW is only statistically significant in the third quintile. Furthermore, looking at

the CMA factor, it is seen that it is only statistically significant in the first and third quintiles. This means that the Fama & French (2015) control variables are statistically significant in combination with the Governance score about half of the time. Thus, these control variables are still said to be important to use. Next, the focus of this paper shifts specifically to the most important variable, the 'cons' meaning the alpha. It can be seen in figure 8a, that the alpha does stay fairly constant between quintile 1 and quintile 3. The alpha is also statistically significantly different from 0 in all three quintiles. The found coefficients are interpreted in the next step of the research, where the differences between quintiles 3 and 1 are being looked at.

This research will continue by looking at the three Governance-created Value-Weighted quintiles in figure 8b. It can be seen that the mkt-RF is the only statistically significant control variable in all three quintiles. For the two other control variables of the Fama & French (1992) paper, it is found that only the first and third quintiles of both the SMB and HML variables are statistically significant. This confirms that the Fama & French (1992) factors are good factors in general and important too, since they explain some variation in at least one of the quintiles. Furthermore, when looking at the Fama & French (2015) factors, it is seen that only quintile 3 of RMW is statistically significant. All the other RMW and CMA control variables are not statistically significant. This is a big deviation from the findings of Fama & French (2015). It seems that these control variables have less explanatory power in combination with the Governance score. Comparing the amount of significant RMW and CMA variables of the first, second, and third hypotheses with the results of hypothesis 4, a big difference is noticed. The exact reason is difficult to find, but it is an important finding that could be given more attention to by a new research paper. The Russell 1000 index should be fairly similar to the NYSE, NASDAQ, and AMEX used in the Fama & French (1992, 2015) paper. Finally, when looking at the 'cons', which is the alpha, it can be seen that the alpha barely moves between the three quintiles, though some small movements happen between quintiles 3 and 1. This is contrary to our expectations beforehand, but is less surprising when considering the given that the control variables also appear rather weak. All three 'cons' or alphas have high t-scores, meaning that they are statistically significantly different from 0.

Name:	EW return q. 1 gov	EW return q. 2 gov	EW return q. 3 gov
Mkt-RF	1.117478***	1.078782***	1.067443***
	(32.168433)	(41.594970)	(36.791682)
SMB	0.486637***	0.400826***	0.336508***
	(7.635569)	(8.423796)	(6.321870)
HML	0.189901***	0.196811***	0.329145***
	(3.022165)	(4.195249)	(6.271802)
RMW	0.001811	0.105872	0.273752***

	(0.019905)	(1.558649)	(3.602636)
CMA	-0.374422***	-0.095598	-0.195093**
	(-3.521507)	(-1.204292)	(-2.196966)
cons	-0.006492***	-0.006369***	-0.006334***
	(-4.657050)	(-6.119866)	(-5.440792)
r2	.9365	.959	.9482
r2 a	.934	.9574	.9462
N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 8a: This figure shows the average equal-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***.

Name:	VW return q. 1 gov	VW return q. 2 gov	VW return q. 3 gov
Mkt-RF	0.982536***	0.984406***	0.980373***
	(37.524426)	(52.827022)	(61.933000)
SMB	0.096631**	0.014490	-0.137931***
	(2.011528)	(0.423839)	(-4.749397)
HML	-0.103425**	-0.030346	-0.064969**
	(-2.183687)	(-0.900294)	(-2.269025)
RMW	-0.090188	0.092778*	0.229350***
	(-1.315148)	(1.901021)	(5.532075)
CMA	-0.117972	0.077800	-0.038961
	(-1.472041)	(1.364070)	(-0.804148)
cons	-0.003389***	-0.005180***	-0.003756***
	(-3.225494)	(-6.926769)	(-5.912180)
r2	.9381	.966	.9728
r2 a	.9357	.9646	.9717

N	132	132	132
t statistics in parentheses			
="* p<0.10	** p<0.05	*** p<0.01"	

Figure 8b: This figure shows the average value-weight coefficients of all the control variables, based on the three quintiles that the ESG score factor is split up by. Mkt-RF is the average market return minus the risk-free rate (based on the 1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. The r2 stands for R-squared (r^2) and is a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model. The r2_a stands for the adjusted R-squared(r^2_a), which is a corrected goodness-of-fit measure for linear models. Finally, N represents the number of months that are taken into consideration in this research. The P-values are noted by a *, where p<0.1 = *, p <0.05 = **, and p<0.01=***.

After looking at the different quintiles for the four hypotheses, the statistical significance of these differences can now be analyzed by running a regression with Newey & West (1987) standard errors. The results will be controlled for lagged autocorrelation, up to six periods back. This causes the results to be slightly weaker since the standard errors are bigger compared to a regression being done without autocorrelation. This means that the p-values become slightly higher. The Newey & West (1987) standard error method is a robust method/estimator which is very accurate when there might be heteroskedasticity and/or autocorrelation. Furthermore, when there is a lagged value in the panel model of an indicator, then this method is very consistent. This will also allow us to make the final conclusions about our hypothesis.

Now that the focus moves to the final results of this research paper, the paper will look to find out if the top portfolio, with the highest scores, did outperform the portfolio with the lowest scores. For this, again the alphas are looked at. The findings will again be analyzed per hypothesis.

Portfolio analysis for hypotheses 1, 2, 3, and 4

The analysis will start by looking at the first column of figure 9a, the equal-weight ESG quintile stock returns difference. Here, it is found that the control variables mkt-RF, SMB, and CMA were all statistically significant at the 5% level. Furthermore, it can be seen that HML and RMW are not statistically significant at the 5% level. Beforehand, all these factors were expected to be statistically significant. It is found that HML and RMW go against our expectations. Further investigation will need to be done to find the causes of these different results. Looking at the most important finding, it can be seen that the difference between the highest alpha of quintile (3) and the lowest alpha of quintile (1) is statistically significant. For the alpha, it is found that it is statistically significant at the 5% significance level, as it is denoted with two asterisks. This finding can be interpreted. We find that on average, the stocks with the highest ESG score, tend to outperform the stocks with the lowest ESG

score, by a 0.2755% higher stock return per month. With a t-statistic of 2.240048, our findings are clearly statistically significant. We can conclude that we should implement a long-short strategy, where we short the companies with a low ESG score and go long on the companies with a high ESG score. This would lead to a statistically signficicant higher stock returns. This means that the hypothesis can be confirmed, which states that a higher ESG score leads to a significantly different stock return. These findings confirm that there is a positive effect on stock returns. This is comparable to the findings in previously discussed research papers, like La Torre et al. (2020).

When the focus is shifted to the value-weight ESG quintile stock returns difference, a different result is found. The first column of figure 9b shows that the control variables mkt-RF, SMB, and CMA and RMW are all statistically significant on the 5% level. This is a relatively strong result for the control variables. Only the HML factor is found to be not statistically significantly different at the 5% level. Based on the literature, all these factors should be statistically significant, so HML is the only result that goes against our expectations. Looking at the most important finding, it can be seen that the difference between the highest alpha of quintile (3) and the lowest alpha of quintile (1) is not statistically significant. This coefficient can thus not be interpreted. With a t-statistic of 1.108878, our findings are pretty far from being statistically significant. We can not conclude that we should implement a long-short strategy, where we short the companies with a low ESG score and go long on the companies with a high ESG score. This means that the first hypothesis can be rejected in the case of the value-weighted portfolio, which states that a higher ESG score leads to a significantly different stock return. These findings do go against some of the findings in previously stated research papers, like La Torre et al. (2020).

Overall, we can conclude from the results that they are mixed for the effects of the ESG scores on the excess stock returns. With one statistically significant positive result for the equal-weighted stock portfolio, but no significant results for the value-weighted portfolio, it is difficult to make a strong conclusion. Furthermore, the Fama & MacBeth (1973) regression also didn't find any significant results. More research could be conducted to gain more certainty, but for this paper the results are leaning toward a positive or potentially no effect on stock returns.

Next, looking at the second column of figure 9a, the equal-weight Environmental quintile stock returns difference, it is found that the control variables mkt-RF, SMB, and CMA are again all statistically significant on the 5% level. Furthermore, it is found that HML is not statistically significant at the 5% level, this time together with the RMW factor. The cause for two of the five Fama & French (1992, 2015) factors to be insignificant is still unknown. Further investigation could be done to find the causes of these different results. When the focus shifts toward the most important finding, it is found that the difference between the highest alpha of quintile (3) and the lowest alpha of quintile (1) is not statistically significant. This means that the found coefficient cannot be interpreted. With a t-statistic of 0.970785, the findings are clearly not statistically significant. This means that the second hypothesis

is rejected in the case of equal-weight stock returns, which states that a higher Environment score leads to a significantly different stock return. The result of this research paper goes against the findings of papers like Derwall et al. (2005), since no sign is found that an Environment score leads to higher excess returns.

Looking at the second column of figure 9b, it is found that the control variables SMB, RMW, and CMA are all statistically significant on the 5% level. Furthermore, it is found that both Mkt-RF and HML are not statistically significant at the 5% level. This result is relatively surprising, especially since the Mkt-RF factor has been statistically significant in all previous cases. For the HML factor, it is more expected since this factor has been statistically insignificant in a lot more tables. Looking at the most important finding, the alpha, it is found that the difference between the alpha of quintile (3) and the alpha of quintile (1) is not statistically significant. This means that the found coefficient cannot be interpreted. With a t-statistic of -0.072660, our findings are clearly not statistically significant. This means that the second hypothesis, which states that a higher Environment score leads to a significantly different stock return, has to be rejected when looking at the value-weight stock returns. This result goes against our previous expectations, since we do not find any sign that an Environment score leads to higher excess returns.

Overall, we can conclude that the results are relatively straightforward when looking at the effects of the Environment scores on the excess stock returns. With no statistically significant result for both the equal-weighted stock portfolio and the value-weight stock portfolio, the conclusion seems to be that there is no effect between these variables. Furthermore, we found no statistically significant results on the Fama & MacBeth (1973) regression. More research could be conducted on the Russell 1000 index, in order to confirm our findings with more certainty.

Then, looking at the third column of figure 9a, the control variables mkt-RF, SMB, HML, and CMA can be seen, which are all statistically significant on the 5% level for the equal weight stock returns. This shows that the Fama & French Three-Factor model (1992) gave us good control variables. Furthermore, looking further, it is found that RMW is not statistically significant at the 5% level. Further investigation will need to be done to find the cause of this result, as it is contrary to the literature reviewed in our research paper. As could be seen in the previous step, the individual quintiles of the RMW factor were statistically significant, so this result is surprising. Focusing on the most important finding, the difference between the highest alpha of quintile (3) and the lowest alpha of quintile (1) is statistically significant at the 5% level. This means that the found coefficient can be interpreted. From this finding it can be interpreted that on average, the stocks with the highest Social score, tend to outperform the stocks with the lowest Social score, by a 0.25570% higher stock return per month. We can conclude that we should implement a long-short strategy, where we short the companies with a low Social score and go long on the companies with a high Social score. This would lead to statistically significant higher stock returns. This means that the third hypothesis can be

confirmed, which states that a higher Social score leads to a statistically different increased stock return. This turns out to be a positive effect according to our results. With a t-statistic of 2.210064, these findings are clearly statistically significant. This finding confirms the results also found in Mănescu (2011), who looked at socially responsible investing (SRI) and focused on the concept of 'eco-efficiency'. This paper stated that eco-efficiency can be thought of as the economic value a company creates relative to the waste it generates.

Looking at the third column of figure 9b, the control variables mkt-RF, SMB, HML, RMW, and CMA are all statistically significant on the 5% level. This is a perfect example of the results that were expected beforehand in our research. When we look at the most important finding, we see that the difference between the highest alpha of quintile (3) and the lowest alpha of quintile (1) is not statistically significant. This means that the found coefficient cannot be interpreted. It can be seen that the 'cons', or alpha has a t-statistic of 1.363432. This shows that the coefficient is pretty far from being significant. We can conclude that we should not implement a long-short strategy, where we short the companies with a low Social score and go long on the companies with a high Social score, since we found no statistical significance. This also means that the third hypothesis has to be rejected for the value-weight results, which states that a higher Social score leads to a statistically different increased stock return.

Overall, we can conclude that the results are relatively mixed when looking at the effects of the Social scores on the excess stock returns. We found no statistically significant results on the Fama & MacBeth (1973) regression. Furthermore, with a statistically significant result for the equal-weighted stock portfolio, but no statistically significant results for the value-weight stock portfolio, the conclusion seems to be that there could be a positive or no relationship between the Social score and the excess stock returns. More research could be conducted on the Russell 1000 index, in order to confirm our findings with more certainty.

Finally, looking at the fourth column of figure 9a, the control variables SMB, HML, RMW, and CMA are all statistically significant on the 5% level in the equal weight return table. The only variable not statistically significant is the Mkt-RF variable. This is surprising, given the fact that this variable was significant in all the previous steps. This variable is insignificant with a t-statistic of 1.946341, which means that the Mkt-RF factor is very close to being statistically significant. Further investigation will need to be done to answer why this variable is not statistically significant. It could be possible that this is just a case of bad luck, and could be appointed to the 5% significance level. The most important finding of this column contradicts the existing literature somewhat. Looking at the alpha, it is found that the difference between the highest quintile (3) and the lowest quintile (1) is not statistically significant. With a t-statistic of 0.133644, the conclusion is that this coefficient cannot be interpreted. Furthermore, this also means that the fourth hypothesis has to be rejected, which states that a higher Governance score leads to a significantly different stock return. This finding still confirms part of the

findings of Bozec & Bozec (2011). In this research paper he states that about 50% of the research papers find a statistically significant effect. In the other 50% of research papers, just like we found in our results, no statistically significant results are found.

Looking at the fourth column of figure 9b, some of the control variables, SMB and RMW are statistically significant at the 5% level in the value-weighted table. It can furthermore be seen that the other variables, Mkt-RF, HML and CMA are all not statistically significant. This is a big difference from the findings of Fama & French (1992, 2015). Also, when we compare these findings to the equal-weighted results, we see that much fewer control variables are statistically significant. This could partly be explained by value-weighted portfolios tending to have weaker results. In this case, there are very few statistically significant control variables when comparing them to the earlier found value-weight results. More research could be conducted to look into the causes of this deviation, since it is not directly evident what causes this. Finally, we see that the most important variable, the 'cons' or alpha, goes against our expectations based on the existing literature. Looking at the alpha, it is found that the difference between the highest quintile (3) and the lowest quintile (1) is not statistically significant. With a t-statistic of -0.208016, this coefficient cannot be interpreted. Furthermore, this also means that the fourth hypothesis has to be rejected, which states that a higher Governance score leads to a statistically different stock return. This finding still partly confirms the findings of Bozec & Bozec (2011). In this research paper, he states that in about 50% of the research paper a statistically significant effect is found. However, in the other 50% of research papers, just like we found in our results, no statistically significant results are found.

Overall, we can conclude that the results are relatively straightforward when looking at the effects of the Governance scores on the excess stock returns. With no statistically significant result for both the equal-weighted stock portfolio and the value-weight stock portfolio, the conclusion seems to be that there is no effect between these variables. When we take into account the findings of the Fama & MacBeth (1973) regression, we see that one negative effect was found. In general, this means that most of our research gathered no statistically significant results, which will be the conclusion of this paper. More research could be conducted on the Russell 1000 index, in order to confirm these findings.

Name:	EW return q. diff. ESG	EW return q. diff. Env	EW return q. diff. Soc	EW return q. diff. Gov
Mkt-RF	-0.088718**	-0.073005**	-0.085597**	-0.050034*
	(-2.121721)	(-2.105875)	(-2.078409)	(-1.946341)
SMB	-0.390140***	-0.375842***	-0.380165***	-0.150129***
	(-6.895336)	(-6.711152)	(-7.295305)	(-3.323925)
HML	-0.031927	0.004775	-0.150857**	0.139244***
	(-0.570440)	(0.069173)	(-2.220445)	(3.815198)

RMW	0.159201	0.157911*	0.067811	0.271941***
	(1.617003)	(1.880364)	(0.783724)	(5.491376)
CMA	0.295068***	0.379298***	0.195547**	0.179329***
	(3.804093)	(4.260353)	(1.985824)	(2.674173)
cons	0.002755**	0.001215	0.002557**	0.000157
	(2.240048)	(0.970785)	(2.210064)	(0.133644)
r2				
r2 a				
N	132	132	132	132
t statistics in parentheses				
="* p<0.10	** p<0.05	*** p<0.01"		

Figure 9a: This figure shows the final results after creating the portfolios by doing a Newey & West (1987) regression on the equal-weight results. We control the ESG, Environment, Social, and Governance -scores for the same Fama & French Five-factor model. Mkt-RF is the average market return subtracted with the Risk-free rate (1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. Finally, N represents the number of months that are taken into consideration in this research.

Name:	VW return q. diff. ESG	VW return q. diff. Env	VW return q. diff. Soc	VW return q. diff. Gov
Mkt-RF	-0.079947***	-0.064410*	-0.082381**	-0.002163
	(-2.891681)	(-1.927457)	(-2.533350)	(-0.071661)
SMB	-0.398477***	-0.388875***	-0.376372***	-0.234562***
	(-6.721223)	(-6.728464)	(-6.549823)	(-4.311338)
HML	-0.018365	0.019496	-0.151459**	0.038455
	(-0.247427)	(0.230998)	(-2.397216)	(0.338989)
RMW	0.340968***	0.356618***	0.348614***	0.319539***
	(3.920598)	(4.205982)	(4.115940)	(3.322679)
CMA	0.487217***	0.555916***	0.392781***	0.079011
	(3.905152)	(4.435906)	(3.440228)	(0.474251)
cons	0.001373	-0.000083	0.001574	-0.000366
	(1.108878)	(-0.072660)	(1.363432)	(-0.208016)
r2				
r2 a				

N	132	132	132	132
t statistics in parentheses				
="* p<0.10	** p<0.05	*** p<0.01"		

Figure 9b: This figure shows the final results after creating the portfolios doing a Newey & West (1987) regression on the Value weighted results. We control the ESG, Environment, Social, and Governance -scores for the same Fama & French Five-factor model. Mkt-RF is the average market return subtracted with the Risk-free rate (1 month Treasury Bill), SMB stands for the Fama & French factor Small Minus Big, likewise HML stands for High Minus Low, RMW stands for Robust Minus Weak and CMA stands for Conservative Minus Aggressive. Finally, N represents the number of months that are taken into consideration in this research.

Limitations & future research

There are several limitations to this research paper, which could be done more extensively. This paper only did a Fama & MacBeth (1973) regression on the Three-Factor Model, but the results of also doing it on the Five-Factor Model could have been interesting as well. This could be looked into by another research paper. Likewise, the results of the portfolio analysis being done separately on the Fama & French (1992) Three-Factors could also have given interesting results. There is relatively little research done on the Russell 1000 index specifically, which means that many research questions could still potentially be tested here. Another weakness of the research is that the Book-to-market ratio was lagged for six months, instead of making it lagged to the last published book-to-market value. The lack of a more thorough explanation for some of the insignificant results, specifically for the control variables is another part of this research paper that could be improved. Further investigation could be done as to why the Russell 1000 has such differing results. When we look at the result of the negative Governance score on the excess stock returns, it is relatively surprising that this result was only found in the Fama & MacBeth (1973) regression. Further investigation could be done to look at the exact effects of the Governance score on excess stock results. The view of Bozec & Bozec (2011) on this matter is still very interesting. They write that the absence of a clear (or even negative) relation between overall governance and performance in North America might be interpreted as a sign that Canadian and US firms have already adopted relatively good governance practices. This makes that any additional improvements in overall governance have no significant impact on their performance, and thus that it is difficult to find any statistically significant results that you would expect beforehand.

Conclusion

So to conclude, the following conclusions can be drawn from the results. Looking at the results of the Fama & MacBeth (1973) regression with the Three-Factor model control variables, it is found that hypotheses 1, 2, and 3 are all rejected at the 5% significance level. For the fourth hypothesis, it is found that the Governance score coefficient is statistically significant. It is found that on average, in our sample, when the Governance score is one point higher, the average monthly stock return becomes -0.0045% lower. The found total difference between the highest and lowest scores is equal to -0.446895% stock return per month.

The results of the portfolio analysis with the Five-Factor model control variables, found that the first hypothesis, where it was stated that a higher ESG score leads to a significantly different excess stock return, was accepted. It was found that on average, the stocks with the highest ESG score, tend to outperform the stocks with the lowest ESG score, by a 0.2755% higher stock return per month in the case of equal-weighted stocks. No statistically significant results for the value-weighted results were found. The second hypothesis, stating that a higher Environment score leads to a significantly different excess stock return, was rejected at the $p<0.05$ significance level for the equal-, and value-weighted returns. The third hypothesis, stating that a higher Social score leads to a significantly different stock return, was accepted for the equal-weight results. It was found that on average, the stocks with the highest Social score tend to outperform the stocks with the lowest Social score by a 0.25570% higher stock return per month. It was found that there are no statistically significant results in the case of value-weighted stocks. Finally, the fourth hypothesis stated that a higher Governance score leads to a significantly different stock return, which was rejected at the $p<0.05$ significance level for both the equal-, and value-weighted returns.

References

Admin. (2022, April 17). *ESG: The report*. ESG The Report. Retrieved May 10, 2022, from https://www.esgthereport.com/what-is-an-ESG score/#Who-calculates-the-score

Akgun, O. T., Mudge, T. J., & Townsend, B. (2021). How company size bias in ESG scores impacts the small cap investor. *The Journal of Impact and ESG Investing*, *1*(4), 31–44. https://doi.org/10.3905/jesg.2021.1.022

Albertini, E. (2013). Does environmental management improve financial performance? A meta-analytical review. *Organization & Environment*, 26(4), 431–457. https://doi.org/10.1177/1086026613510301

Ambrose, J. (2021, May 26). *ExxonMobil and Chevron suffer shareholder rebellions over climate.* The Guardian. Retrieved March 26, 2022, from https://www.theguardian.com/business/2021/may/26/exxonmobil-and-chevron-braced-for-showdown-over-climate

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. https://doi.org/10.2469/faj.v74.n3.2

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/10.1016/0304-405x(81)90018-0

Bauer, R., Guenster, N., & Otten, R. (2004). Empirical evidence on corporate governance in Europe: The effect on stock returns, firm value and performance. *Journal of Asset Management*, 5(2), 91–104. https://doi.org/10.1057/palgrave.jam.2240131

Bozec, R., & Bozec, Y. (2011). The use of governance indexes in the governance-performance relationship literature: International evidence. *Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences De L'Administration*, *29*(1), 79–98. https://doi.org/10.1002/cjas.201

Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate Social Performance and Stock Returns: UK evidence from Disaggregate Measures. *Financial Management*, 35(3), 97–116. https://doi.org/10.1111/j.1755-053x.2006.tb00149.x

Chan, L. K., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *The Journal of Finance*, *46*(5), 1739–1764. https://doi.org/10.1111/j.1540-6261.1991.tb04642.x

Cornell, B. (2020). ESG preferences, risk and return. *European Financial Management*, 27(1), 12–19. https://doi.org/10.1111/eufm.12295

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal*, *61*(2), 51–63. Https://doi.org/10.2469/faj.v61.n2.2716

Dirkx, P., & Peter, F. J. (2020). The Fama-French five-factor model plus momentum:

Evidence for the German market. *Schmalenbach Business Review*, *72*(4), 661–684. https://doi.org/10.1007/s41464-020-00105-y

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Freedman, M., & Patten, D. M. (2004). Evidence on the pernicious effect of financial report environmental disclosure. *Accounting Forum*, 28(1), 27–41. https://doi.org/10.1016/j.accfor.2004.04.006

Ganti, A. (2022, February 8). *Russell 1000 index*. Investopedia. Retrieved May 12, 2022, from https://www.investopedia.com/terms/r/russell_1000index.asp

Hulsebosch, Shelley. (2020, July 16). The ESG risk factor in equity pricing models: A study between the United States and Europe. *Finance & Investments*. Retrieved from http://hdl.handle.net/2105/54348

La Torre, M., Mango, F., Cafaro, A., & Leo, S. (2020). Does the ESG index affect stock return? evidence from the eurostoxx50. *Sustainability*, 12(16), 6387. https://doi.org/10.3390/su12166387

Lintner, J. (1975). Inflation and security returns*. *The Journal of Finance*, *30*(2), 259–280. https://doi.org/10.1111/j.1540-6261.1975.tb01809.x

Mănescu, C. (2011). Stock returns in relation to environmental, social and governance performance: Mispricing or compensation for risk? *Sustainable Development*, 19(2), 95–118. https://doi.org/10.1002/sd.510

Newey, W. K., & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, *28*(3), 777. https://doi.org/10.2307/2526578

Nollet, J., Filis, G., & Mitrokostas, E. (2016). Corporate Social Responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, 400–407. https://doi.org/10.1016/j.econmod.2015.09.019

Rehnby, N. & Karlstad Business School. (2014). *Does the Fama-French three-factor model and Carhart four-factor model explain portfolio returns better than CAPM? A study performed on the Swedish stock market.* Retrieved from https://www.diva-portal.org/smash/get/diva2:944713/FULLTEXT04.pdf

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Thune, K. (2022, March 9). *Should you invest in a total stock index fund or S&P 500 index fund?* The

Balance. Retrieved May 12, 2022, from https://www.thebalance.com/total-stock-market-vs-sandp-500-2466403

Velte, P. (2017). Does ESG performance have an impact on financial performance? evidence from Germany. *Journal of Global Responsibility*, 8(2), 169–178. https://doi.org/10.1108/jgr-11-2016-0029