UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Navigating in the ESG score jungle

A cross-sectional approach to determine the ESG risk factor

Gustav Pettersson* Mattias Öhrn†

June 2022

A thesis presented for the degree of
Master of Science in Finance

Thesis Advisor: Evert Carlsson
The School of Business, Economics and Law
Graduate School

*guspetguh@student.gu.se
†gusmathr@student.gu.se

Abstract

This thesis examines the relationship between ESG scores and yearly excess return between 2010 and 2020 on the S&P 500 Index. With a solid theoretical background regarding investor preferences, we ask whether investors accept lower returns for holding greener assets. Our method is a cross-sectional approach, using pooled time-series regressions and Fama-MacBeth regressions, where we seek to determine the ESG risk score factor. We find significant evidence that ESG scores have a negative relationship with yearly excess return in all our regressions when controlling for other return predictors and the *Sin Stock anomaly.* This relationship holds for the overall ESG score and the separate ESG pillar scores, Environmental, Social, and Governance. Our results prove to be consistent with previous research regarding ESG-motivated investors. We found inconsistent results with previous research regarding the Governance pillar score, arguing that the Governance pillar score may not be an appropriate proxy. Our results remain consistent while conducting further robustness tests with clustering on the sector level.

Keywords: Asset Pricing, ESG Investing, ESG Risk Score, Factor Models, Fama-MacBeth Regressions, Time-Series Regressions.

JEL Classifications: G11; G12; M14.

Acknowledgements

First and foremost, we would like to express our special thanks and gratitude to our supervisor, Evert Carlsson, who enabled this thesis with his guidance and support. We would also like to thank the School of Business, Economics and Law, and the Centre for Finance for these academic years.

Contents

List of Tables

1 Introduction

This thesis investigates the relationship between ESG scores and excess return by including a theoretical framework where assets become consumer goods (Daniel and Titman, 1997, Fama and French, 2007). Hence investors' utility functions differ from traditional mean-variance decomposition. This framework brings new questions regarding how investors should incorporate ESG in their investment decisions. Previous studies have focused on ESG investing as an investment strategy and if investors holding greener stocks outperform the market (Halbritter and Dorfleitner, 2015, Naffa and Fain, 2022). Our approach examines whether investors in the US market are willing to accept a lower premium for holding stocks with higher ESG scores. When investors' utility functions include their preferences for ESG, the price of assets should be affected (Pedersen *et al.*, 2021).

We apply this framework on the S&P 500 Index spanning over ten years, where we run Fama-MacBeth regressions and pooled time-series regressions controlling for other return predictors such as *Size* (Banz, 1981) and *Value* (Fama and French, 1993). We contribute to the research field of ESG investing through our cross-sectional approach and by controlling for the *Sin Stock anomaly* (Hong and Kacperczyk, 2009). Furthermore, we analyze the relationship between the separately ESG pillar scores and excess returns to explore whether these pillar scores' relationship between returns differs.

Daniel and Titman (1997) question whether firm characteristics or *factor loadings* can explain return in their characteristic model. In effect, they assume preferences for assets as consumption goods. Fama and French (2007) extend this model when analyzing how tastes for assets can affect prices. By allowing investors to have preferences for assets, the investors gain the utility of holding certain assets, thus influencing prices.

In recent studies, Pástor *et al.* (2021) and Pedersen *et al.* (2021) use the foundation of Fama and French (2007) by including investors' tastes for ESG; they show how different investors depending on their preference for ESG, have separate conditional expected returns; thus, the equity risk premium for holding green assets decreases. While Pástor *et al.* (2021) develop their own ESG risk score, we follow the structure of Pedersen *et al.* (2021), which uses proxies for ESG to explain excess return. We apply ESG scores from Refinitiv Eikon as proxies for ESG and historically proven return predictors to perform our pooled time-series regressions and Fama-MacBeth regressions.

We establish the idea of ESG scores as proxies for investors' preferences regarding environmental, social, and governance issues. However, there is a dispute regarding the significant relationship between ESG scores as a risk factor and excess return. For example, Maiti (2021) finds a relationship where ESG scores predict returns, and Friede *et al.* (2015), which investigates over 2000 empirical studies on ESG and financial performance, support this thesis. On the other hand, Halbritter and Dorfleitner (2015) question this relationship and argues that the magnitude and direction hardly depend on the rating provider, period, and companies included.

Besides ESG scores, another shown relationship is between firms' carbon emission and excess return (Bolton and Kacperczyk, 2021), a relationship where higher carbon emissions indicated a higher excess return. Hence investors demanded a premium for holding these brown assets. One can compare these results to the *Sin Stock anomaly* in Hong and Kacperczyk (2009), where they found that investors, to some extent, demanded a premium for holding unethical assets.

One can state a discrepancy in this research field where ESG scores seem to predict returns in one way or another, and other environmental and social proxies, on the other hand, shows a clear investment premium for assets with high exposure to carbon emissions and unethical businesses. Besides, according to Pástor *et al.* (2021) and Pedersen *et al.* (2021), ESG-motivated investors should accept a lower premium for holding green assets due to their preferences. We act on this discrepancy where we include both ESG scores and separately pillar scores for Environmental, Social, and Governance to find supporting arguments for the theory of Pedersen *et al.* (2021).

We find evidence that ESG scores have a negative relationship with yearly excess return at a 0.1% significance level. This effect holds even when controlling for the *Sin Stock anomaly.* Our cross-sectional approach yields the same results with pooled time-series regressions and Fama-MacBeth regressions. These results align with the theory of Pedersen *et al.* (2021) by having similar negative coefficients; thus, ESG-motivated investors can accept lower returns due to their utility function. Having said that, investors demand a premium for holding stocks with low ESG scores.

The relationships are negative and significant when examining the separately ESG pillar scores. This result is in line with Bolton and Kacperczyk (2021), who argue for the *Carbon Emission premium*, and in line with Pedersen *et al.* (2021) regarding social issues. According to theory and previous research, the Governance factor has been positive or zero (Bebchuk *et al.*, 2013) compared to our results. One possible reason for our different results is that the Governance score pillar does not work as a proxy for good governance.

We present previous literature in Section 2 regarding ESG scores and ESG investing. Section 3 introduces the theoretical framework of investor preferences and shows how different expected returns depend on their taste for assets. In Section 4, we present our data and calculations. The methodology is in Section 5, where we establish our model specifications and the regression methods. Section 6 shows the results from our regressions together with robustness tests. Section 7 states our contributions, limitations, and further research suggestions.

2 Literature Review

Various researchers in the field of asset pricing have found and suggested risk factors that explain excess return. In earlier studies; we have, for example, *Size* (Banz, 1981), *Value* (Fama and French, 1993), *Momentum* (Jegadeesh and Titman, 1993, Carhart, 1997), and *Profitability & Investment* (Fama and French, 2015). These risk factors are used frequently in factor models and as return predictors, while more recent researchers have investigated whether to include a new ESG risk factor and if it can explain excess return (Friede *et al.*, 2015, Maiti, 2021, Naffa and Fain, 2022). We propose a cross-sectional approach for estimating the ESG risk score using Fama-MacBeth regressions and pooled time-series regressions on the S&P 500 Index. Based on the availability of ESG data and previous literature, we use a time period spanning over ten years. We control for the proxies from the Fama-French five-factor model and use ESG scores as proxies for ESG. The following section will now present other previous studies and their findings.

Using portfolio sorting similar to Fama and French (1993), Maiti (2021) finds that ESG factors predict returns on the STOXX Europe 600 Index. With a three-factor model with market, size, and ESG, the model outperforms the traditional Fama-French three-factor model. Maiti (2021) concludes that ESG factors should not be ignored by investors when making investment decisions and uses the Bloomberg ESG scores as proxies. He comments his findings as mixed, that earlier studies, for example, Friede *et al.* (2015), find a positive relationship between ESG and financial performance and, on the other hand, argue studies like Revelli and Viviani (2015) that the previous ESG studies are inclusive and ambiguous. Maiti (2021) also comments on the problems of finding ESG data and the definition and qualification of the ESG variable.

With Fama-MacBeth regressions and by constructing ESG portfolios sorted on firms' ESG performance from leaders to laggards, Naffa and Fain (2022) found, in contrast, no evidence of significant alphas for ESG portfolios during 2015-2019. Unlike Maiti (2021), do Naffa and Fain (2022) extend the Fama-French five-factor model with new ESG factors instead of the Fama-French three-factor model. Hübel and Scholz (2020) use a Fama-French five-factor model complemented with the momentum factor to control for other risk factors. They split the ESG factor into the three separate ESG pillar scores' factors, including environmental, social, and governance, making their model a nine-factor model. Hübel and Scholz (2020) find that when including the ESG risk factors, the explanatory power of classical asset pricing models increases. Important to notice is the rejection framework from Harvey *et al.* (2016), who argue that to include another explanatory factor, and avoid data mining, one should not accept t-statistics less than 3.0.

Similar to Maiti (2021), Halbritter and Dorfleitner (2015) investigated the link between financial performance and ESG ratings. Their study used ESG data from ASSET4 (now Refinitiv Eikon), Bloomberg, and KLD and stock data from the US market between 1991 to 2012. The econometrical framework includes both ESG portfolios using the Carhart four-factor model (Carhart, 1997) and a cross-sectional approach using Fama-MacBeth regressions. Their main findings were that investors no longer should expect higher returns by trading portfolios concerning ESG aspects. Additionally, the Fama-MacBeth regressions show that the magnitude and direction of impact are mainly dependent on the rating provider, the company, and the time period (Halbritter and Dorfleitner, 2015).

Another approach presented by Pástor *et al.* (2021) is their two-factor model. Their research shows that investing in green assets considering ESG criteria in equilibrium gains lower expected returns since investors tend to hold them for other reasons than expected returns. On the other hand, green assets tend to outperform when positive shocks linked to their ESG factor happen.

Pedersen *et al.* (2021) propose a theory where the ESG scores contribution is twofold. First, it provides information about the firm's fundamentals, and second, it affects the investors' preferences. As a proxy for ESG, they use the MSCI ESG score. For the separate E, S, and G, they use carbon emissions (E), Sin stock (S), and accruals (G) and find no or weak evidence for an ESG and carbon emissions premium. They found evidence of a sin premium, whereas the governance premium shows substantial significant results that investors should consider. Like Pástor *et al.* (2021), do Pedersen *et al.* (2021) investigate asset prices in equilibrium where they derive an ESG-efficient frontier that shows the highest Sharpe ratios for each respecting ESG level.

While Pedersen *et al.* (2021) studied both ESG scores and E, S, and G separately, Bolton and Kacperczyk (2021) focused on the Environmental pillar and explored the impact of carbon emissions on stock returns of US companies. They argued that investors demand a risk premium when investing in companies with high carbon emissions. This risk premium exists because of the ethical reasons for investing in these companies and their exposure to new carbon laws, such as an implemented emissions tax that would negatively shock the firms.

Bolton and Kacperczyk (2021) use a cross-sectional analysis and start by posing three hypotheses. The *Carbon risk premium hypothesis*, because there is a correlation between high levels of emissions and energy use of fossil fuels, thus corporate returns are affected by price controls on fossil fuels and raw materials. The *Market inefficiency hypothesis*, as exposure to regulation, may be a risk factor at the firm level; Simultaneously, regulating a federal carbon tax could be done at the industry level. It tests whether financial markets underestimate the risk associated with carbon emissions by controlling for known risk factors, industry, and firm characteristics. The *Divestment hypothesis* means that firms with high carbon emissions are not different from other kinds of Sin stocks. Institutional investors refrain from investing in these firms due to social responsibility and ethical reasons, leading to higher stock returns; they question whether investors are looking at the company level or whether the breakdown is broader, such as at the industry level.

Using a comprehensive data set of seven data providers, Bolton and Kacperczyk (2021) attempt to cover the entire US universe of companies with the emissions data available. Similar to Maiti (2021) and Halbritter and Dorfleitner (2015), they use Refinitiv Eikon and Bloomberg but supplement it with data from CDP, Trucost, MSCI, Sustainalytics, and ISS; as a result, their sample includes 3421 companies from 2005 to 2017. They find a positive and statistically significant relationship between carbon emission and stock return by running pooled fixed effects regressions, similar to Pedersen *et al.* (2021). Statistically, the significant result was confirmed in an industry fixed effects regression, as emissions tend to cluster at the industry level, according to Bolton and Kacperczyk (2021).

One can compare Pástor *et al.* (2021) results considering that green assets gain lower expected returns with Hong and Kacperczyk (2009), who investigated the performance of Sin stocks on the US stock market as well as the global stock market. Their findings were that Sin stocks tended to outperform the market between 1965 and 2006. Hong and Kacperczyk (2009) used the classification of Sin stocks as in (Fama and French, 1997); hence, *Beer & Alcohol*, *Smoke & Tobacco*, and *Gaming*. The idea of Sin stocks is that investors demand a higher expected return holding these compared to green assets (Hong and Kacperczyk, 2009).

While Bolton and Kacperczyk (2021) investigated the relationship between carbon emission and return and Hong and Kacperczyk (2009) investigated the *Sin Stock anomaly*, did Bebchuk *et al.* (2013), in contrast, examine the relationship between corporate governance and return. Bebchuk *et al.* (2013) found that the earlier documented correlation between good governance and the abnormal return has disappeared. They argue that this disappearance is that investors gradually learned the usefulness of governance and included it in their investment decision; hence it is already incorporated in the price of assets. On the other hand, Pedersen *et al.* (2021) find that their proxy for governance (low accruals) correlate with future returns. This relationship was true for several factor models, including the Capital Asset Pricing model (Sharpe, 1964, Litner, 1965, Mossin, 1966); the Fama-French three-factor model (Fama and French, 1993); and the Fama-French five-factor model (Fama and French, 2015).

3 Theory

3.1 Factor Pricing Models

As Markowitz (1952) built the cornerstone for asset price modeling, the Capital Asset Pricing model by Sharpe (1964), Litner (1965), and Mossin (1966) has expanded the mean-variance-efficiency model, that investors only care about expected returns and the return dispersion. Indeed, the founding of the CAPM contributes to asset price modeling being a straightforward model, but today's knowledge violates the weak assumptions it comes with (DeMarzo and Berk, 2016). For instance, Kahneman and Tversky (1979) confronted rational behavior and financial decision-making and found evidence that people cannot be homogeneous and rational in decision-making as individuals react differently to gains and losses. Furthermore, the same information will be interpreted differently by individuals. This evidence also leads to violations of the general theory, the *Efficient market hypothesis*, that all stock prices consider all available information, and thus the market cannot be beaten (Kahneman and Tversky, 1979).

The Nobel laureate Eugene Fama and Kenneth French introduced their five-factor model in 2015 to determine asset pricing, as demonstrated in Equation 1. The five-factor model uses four additional factors to estimate expected returns; a proxy for size *(SMB)*, value *(HML)*, profitability *(RMW)*, and investment *(CMA)*.

$$R_i - R^f = \alpha_i + \beta_1(E[R_m] - R^f) + \beta_2 SMB + \beta_3 HML + \beta_4 RMW + \beta_5 CMA + \epsilon_i \quad (1)$$

The size factor *SMB* stands for small minus big and represents the difference in returns between portfolios with small-capitalization and large-capitalization, respectively. The value factor *HML* stands for high minus low and explains the difference between portfolios with high book-to-market ratios and low book-to-market ratios, respectively. Their work concluded that one could exhibit excess return higher than the market by shorting stocks with significant capital, whereas they buy stocks with less capitalization and invest in stocks with higher book-to-market ratios relative to the market. Thus, the losses in the short run become rewarded by profits in the long run (Fama and French, 1993). Robust minus weak (RMW) helps explain the expected returns due to diversified portfolios associated with high profitability. Conservative minus aggressive (CMA) focuses on the internal investments of the firms and assumes firms that invest in large growth projects make losses on the stock market (Fama and French, 2015).

3.2 Equilibrium Asset Pricing

The common assumption in asset pricing theory is that agents only care about the expected payoff; thus, investment assets are not consumption goods (Fama and French, 2007). On the other hand, investors have different preferences and utility functions when analyzing assets as consumption goods. One can consider two cases according to Fama and French (2007): (i) investors' utility depends directly on the quantities of assets held, and (ii) the tastes for assets are related to the covariances of asset returns with common return factors or state variables.

Fama and French (2007) introduce two different groups of investors, group A and group D, which evaluate assets differently. Group A evaluates assets on their expected payoffs, similar to common asset pricing theory. This group has no specific tastes for stocks as consumption goods. Hence their utility function does not depend on holding any assets, and the utility function for investor i is then $U_i(C_1, W_2)$ where $W_2 = \Sigma_j q_j(1 + R_j)$. Where C_1 is consumption today and W_2 is wealth tomorrow, the investors' wealth tomorrow depends on the assets return. The proposed investors that gain utility of holding certain assets are called group D, and their utility function is instead $U_i(C_1, q_1, ..., q_j, W_2)$ where $q_1, ..., q_j$ are assets concerned as consumption goods. Here, investors choose assets based on their preferences with possible lower return R_j as a trade-off. The utility of holding different assets for investors in group D can differ.

Pedersen *et al.* (2021) show that an ESG-adjusted CAPM gives the equilibrium security prices and returns. In their theoretical framework, there instead exist three types of investors, type-U (ESG-unaware), type-A (ESG-aware), and type-M (ESG-motivated). The ESG-CAPM then shows that stocks with high ESG scores have high expected returns if there are many type-U investors and when high ESG predicts future earnings. The reason is that the stocks in this market are profitable, and the type-U investors have not bid up the prices of ESG stocks.

If the economy instead has a lot of type-A investors, these investors then bid up the price of the profitable ESG stocks to reflect their expected future returns (Pedersen *et al.*, 2021). Hence, the relationship between ESG and expected returns disappears. Moreover, with instead many type-M investors, the expected return of high-ESG stocks decreases. Since ESG-motivated investors can accept a lower return for holding a higher ESG portfolio, it holds.

The thesis by Pástor *et al.* (2021) supports that investors enjoy holding green assets and thus can accept lower expected returns. While looking at the market instead of the sole investor, Pástor *et al.* (2021) show how the equity premium depends on the average of ESG tastes and the market portfolio's overall "greenness". Thus the market risk premium, according to Pástor *et al.* (2021), is determined by

$$\mu_m = a\sigma_m^2 - \frac{\bar{d}}{a} w_m' g. \tag{2}$$

Where σ_m^2 is the variance of the market return, and a is the investor's relative risk aversion. The parameter $\bar{d}$ is the average of ESG tastes which Pástor *et al.* (2021) multiply with the overall "greenness" of the market portfolio $w_m'g$. Thus, the equity premium will decrease in a net green market, i.e., $w_m'g > 0$ together with a strong taste for ESG, $\bar{d}$ being positive and large. On the other hand, when $w_m'g < 0$ but $\bar{d}$ is the same, the equity premium will increase. Thus, investors with a higher taste for ESG would demand a higher premium for holding brown assets. In an ESG neutral market, we have $w_m'g = 0$; hence the market risk premium only depends on the variance of market return and the investor's relative risk aversion (Pástor *et al.*, 2021).

3.3 ESG-Efficient Frontier

Derived from Pedersen *et al.* (2021), an investor can choose a portfolio of n risky assets and a risk-free bond in the investment universe. The bond denotes r^f, and the securities excess returns as a vector $r = (r^1, ..., r^n)'$. All assets also have ESG scores denoted as $s = (s^1, ..., s^n)'$. Pedersen *et al.* (2021) present three different types of investors and how the market equilibrium is derived.

The first investor is ESG-unaware. This investor does not incorporate the firm's ESG scores in decision-making; thus, the standard CAPM equilibrium exists. For investors of type-U, the unconditional expected excess return is then

$$E\left(r_t^i\right) = \beta^i E(r_t^m). \tag{3}$$

When the investor of type-U ignores ESG, the investor will hold the portfolio, that maximizes the Sharpe ratio. If the investor instead is ESG-aware, meaning that it uses the ESG scores to understand risk and expected return, $\mu = E(r|s)$, its expected return is conditional on ESG information. This relationship gives us the formula for conditional expected returns

$$E\left(r_t^i|s\right) = \beta^i E(r_t^m) + \lambda \frac{s^i - s^m}{p^i}. \tag{4}$$

Given $\lambda > 0$, which means that a high ESG score implies high expected future earnings and that a firm's ESG score (s^i) is higher than the market (s^m), the conditional expected excess return increases for an ESG-aware investor. Recall that if only ESG-aware investors exist in the market, we have conditional CAPM equilibrium. The possibility for the investor to profit by using the information value of the ESG score (λ) disappears since the information is already incorporated in the price (p^i) (Pedersen *et al.*, 2021).

On the other hand, when the market instead consists of ESG-motivated investors with preferences for a high ESG score portfolio. The conditional expected excess return decreases, which Equation 5 denotes

$$E\big(r_t^i|s\big) = \bar{\beta}^i E(r_t^m|s) - \pi\big(s^i - s^m\big). \tag{5}$$

Where π is the scaling parameter of ESG preferences, recall that this parameter only exists for ESG-motivated investors'. The conditional expected return decreases when having strong preferences for ESG, thus π positive and large, and choosing stocks with an average ESG score above the market $(s^i - s^m)$. Additionally, we are back in traditional mean-variance optimization by having investors with no preference for ESG, i.e. $(\pi = 0)$.

3.4 Hypotheses Development

Using the theoretical framework of Pedersen *et al.* (2021) and Pástor *et al.* (2021), which shows that ESG-motivated investors can accept a lower premium for holding assets with high ESG scores, we test if a negative relationship between yearly excess return and ESG scores exists. Furthermore, the previous literature from Bolton and Kacperczyk (2021), Hong and Kacperczyk (2009), and Bebchuk *et al.* (2013) lead us to test the relationships between yearly excess return and Environmental, Social, and Governance pillar scores separately.

Derived from literature and theory, we propose three hypotheses. The later subsection, Model Specification, illustrates a more detailed interpretation and explanation of the models we perform. Our first hypothesis is that the investors are willing to accept a lower equity premium for holding green assets using the theory of Pedersen *et al.* (2021) for ESG-motivated investors. Thus, a higher ESG score will be associated with lower excess return. Therefore, the first hypothesis states that *The relationship of the ESG risk score to yearly excess return is negative and significant.*

Our second hypothesis investigates the relationship between ESG score and the *Sin Stock anomaly* (Hong and Kacperczyk, 2009). Similar to Pedersen *et al.* (2021), the idea is that Sin stocks have positive abnormal returns; hence controlling for them, the ESG coefficient will be more negative compared to Hypothesis 1. We formulate Hypothesis 2 as *The ESG risk score has a more extensive significant negative relationship to yearly excess return when controlling for the Sin Stock anomaly.*

We find support for testing the ESG scores separately as the ESG pillar scores have shown different relationships with excess return (Pedersen *et al.*, 2021, Hübel and Scholz, 2020). This method can eliminate the risk of the overall variable being averaged to a mean far away from their true effect if one variable is positive whereas the other is negative. Similar to Hypothesis 1, we expect the coefficient for *Environmental* to be negative as Bolton and Kacperczyk (2021) showed a significant positive relationship between carbon emissions and return. We also expect the coefficient for *Social* to be negative (Hong and Kacperczyk, 2009). Lastly, we expect a positive or weak coefficient for *Governance* (Bebchuk *et al.*, 2013). By allowing different coefficients for the scores, we formulate Hypothesis 3 as *The relationships between yearly excess return and the separate ESG pillar scores will be significant. The Environmental and Social pillar scores will be negative, and the Governance pillar score will be weak positive.*

4 Data

4.1 Data Description

This study uses the S&P 500 Index, which consists of listed shares on the New York Stock Exchange (NYSE) and NASDAQ in the United States—collecting data from January 2010 to December 2020 yearly covering 505 constituents. WRDS (Wharton Research Data Services) and S&P Capital IQ are the primary sources for stock prices and firm characteristics. The risk-free rate is from Kenneth French's Data Library, and Environmental, Social, Governance, and ESG scores are from Refinitiv Eikon. As the ESG coverage decreases when extending the time period further, together with previous studies using a similar time period, our sample period becomes ten years.

After standardizing ticker symbols between sources, we used the statistical software R to structure and merge data sets using date and ticker as our primary identifier. After matching the data sets, 35 firms missed vital information such as firm characteristics and ESG scores. Therefore, 35 companies were dropped from the data set due to missing values, leaving us with a data set covering 470 listed companies over ten years. Furthermore, we only include firms with information from at least two time periods in the regression models. Merging the data sets by dates and tickers, including firm mergers as the firm retained its ticker, eliminated part of the survivorship bias. However, we do not include discarded or insolvent firms in the study as Halbritter and Dorfleitner (2015). Thus, our study does not eliminate the survivorship bias to the same extent.

4.2 ESG Scores

We use one common provider of ESG data, Refinitiv Eikon, which the financial literature uses as one primary source (Halbritter and Dorfleitner, 2015, Hübel and Scholz, 2020). Refinitiv Eikon covers over 70 percent of the global market and captures over 500 company-level ESG measures (Refinitiv, 2021). Their ESG score ranges from 0.1

for companies with the lowest ESG score to 100 for every data point. Refintiv Eikon weights data points due to importance by dividing the score into ten categories (Refinitiv, 2021), all ten categories and their initial weights are in Table 5 in the Appendix. The weights are then re-balanced for each industry group to target industries' most relevant areas regarding ESG. The relative category weight of the industry is calculated by taking the median value of an industry group and dividing it by the sum of the medians in all industry groups.

Hübel and Scholz (2020) argue that examining the ESG pillars scores separately instead of looking at the aggregated score is the correct test method. The reason is that the aggregated score becomes the mean of all three pillars; hence a positive shock in the Environmental pillar score and a negative shock in the Governance pillar score offset each other. Due to that, the ESG score becomes the mean even since E och G is not close to the mean and strengthens the reason for including the pillar scores E, S, and G separately in this study.

4.3 Financial Variables

In the context of Fama and French (2015), which uses several firm characteristics to create portfolios estimating excess return, our approach differs in the way we collect and structure these characteristics. We do not use portfolio sorting and instead retain all information by running cross-sectional regressions (Galema *et al.*, 2008). The firm characteristics Beta, Size, Book-to-Market, Investments, and Profitability is extracted and calculated with the help of S&P Capital IQ.

The variable *BETA* is the market beta, calculated over five years. More precisely, it is the standard deviation for any given security i due to market risk relative to the market. *LSIZE* refers to the natural logarithm of the firm's market capitalization, similar to Pedersen *et al.* (2021). We calculate *B/M* by dividing the firm's book value of common shareholders' equity by its market capitalization. The variable *INV* is the same as in Bolton and Kacperczyk (2021), where we divide the firm's capital expenditures by its book value of assets. *PROF* is the firm's operating profitability divided by its total assets. Creating a ratio for *PROF* will fix the high skewness and kurtosis in the operating profitability; by conducting a ratio, we get a more normal distributed variable. We winzorise the financial variables at the 2.5% level to avoid outliers and potential value errors without removing important information, as in Bolton and Kacperczyk (2021).

4.4 Descriptive Statistics

We present the descriptive statistics from our independent variables in Table 1. We use yearly observations since Refinitiv Eikon does not update its ESG scores more frequently. Furthermore, some financial variables are from the firm's financial statements; thus, yearly observations make sense. The financial variable *PROF* has slightly higher kurtosis. Otherwise, the distribution of the variables is fine. In total, our data set includes 4568 observations.

Table 1: Descriptive statistics for independent variables

	BETA	LSIZE	B/M	PROF	INV	ESG	E	S	G
Median	1.01	9.67	0.28	0.06	0.03	58.28	53.3	60.17	61.36
Mean	1.02	9.76	0.35	0.08	0.04	56.23	49.0	58.61	58.90
Std. Deviation	0.44	1.14	0.28	0.07	0.03	18.66	28.2	20.72	20.70
Min	0.12	7.08	-0.03	-0.43	0.00	2.46	0.0	1.22	0.45
Max	1.98	12.29	1.18	0.50	0.14	94.62	98.5	98.00	99.10
Skewness	0.09	0.12	1.12	1.13	1.17	-0.38	-0.32	-0.26	-0.41
Kurtosis	2.62	2.97	3.86	6.74	3.83	2.46	1.92	2.27	2.46

Note: This table presents the median, mean, standard deviation, maximum and minimum values, skewness and kurtosis. The data consist of the entire data set of firms from the S&P 500 Index. BETA is the market beta; LSIZE is the natural logarithm of the firms' market capitalization; B/M is the firms' book-to-market; PROF is the firms' operating profitability divided by its total assets; INV is the firms' CAPEX divided by its total assets; ESG is the ESG score from Refinitiv Eikon; E, S, and G are the separate pillar scores.

5 Methodology

5.1 Model Specification

To test our hypotheses, we specify five different time-series models. On the left-hand side, the dependent variable is the yearly excess return for firm i in time period t. In our sample $t = 1, ..., 10$ for yearly return and $i = 1, ..., 470$ for numbers of firms. Hence the return function is as follows $R_{i,t} = \frac{P_{i,t}-P_{i,t-1}}{P_{i,t-1}}$; where $P_{i,t}$ equals the price for asset i in period t. We then subtract the yearly risk-free rate R_t^f to get the yearly excess return. This section starts to estimate Equation 6 to answer Hypothesis 1.

$$R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 ESG_{i,t} + \epsilon_{i,t} \quad \forall \quad t = 1, 2, ..., T \tag{6}$$

Where $R_{i,t} - R_t^f$ is the yearly excess return for firm i in period t; $\boldsymbol{X}_{i,t}$ is a vector that includes the earlier described firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and INV; $ESG_{i,t}$ is the ESG score for firm i in period t and $\epsilon_{i,t}$ is the normally distributed measurement error with mean zero. The argumentation for including vector $\boldsymbol{X}_{i,t}$ is that the Fama-French variables are proven return predictors (Fama and French, 2015) and by controlling for them, increase robust results. This method is similar to Pedersen *et al.* (2021) and Bolton and Kacperczyk (2021).

To control for the *Sin Stock anomaly* (Hong and Kacperczyk, 2009), we include the dummy variable SIN, which consists of 18 firms, and estimates the following regression specification in Equation 7 to answer Hypothesis 2.

$$R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 ESG_{i,t} + \theta_2 SIN_{i,t} + \epsilon_{i,t} \quad \forall \quad t = 1, 2, ..., T \tag{7}$$

$R_{i,t} - R_t^f$ is the yearly excess return for firm i in period t; $\boldsymbol{X}_{i,t}$ is the same vector of firm characteristics, and $ESG_{i,t}$ is the firms' ESG scores as in Equation 6. $SIN_{i,t}$ equals one if the stock is a Sin stock and zero otherwise; $\epsilon_{i,t}$ is the normally distributed measurement error with mean zero. We use the classification as in Fama and French

(1997) and include the following industries *Beer & Alcohol*, *Smoke & Tobacco*, and *Gaming*. We filter on the North American Industry Classification System (NAICS) and the Global Industry Classification Standard (GICS).

Similar to Pedersen *et al.* (2021), we run regressions with the independent variables E, S, and G separately. The method is also in line with Hübel and Scholz (2020), arguing that including the pillar scores separately will increase the statistical power. The reason to run these regressions separately is connected to Hypotheses 3 and tests whether there is a reason to use three different risk factors. Recall that we expect different signs for these coefficients based on previous literature. These regressions are in Equation 8, 9, and 10.

$$R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \theta_1 E_{i,t} + \epsilon_{i,t} \quad \forall \quad t = 1, 2, ..., T \tag{8}$$

$$R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \theta_1 S_{i,t} + \epsilon_{i,t} \quad \forall \quad t = 1, 2, ..., T \tag{9}$$

$$R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma}\boldsymbol{X}_{i,t} + \theta_1 G_{i,t} + \epsilon_{i,t} \quad \forall \quad t = 1, 2, ..., T \tag{10}$$

Where $R_{i,t} - R_t^f$ is the yearly excess return for firm i in period t; $\boldsymbol{X}_{i,t}$ is the same vector of firm characteristics as in Equation 6 and 7. $E_{i,t}$, $S_{i,t}$, and $G_{i,t}$ are the separately conducted pillar scores from Refinitiv Eikon, and $\epsilon_{i,t}$ is the normally distributed measurement error with mean zero.

These models will be compared against each other and tested with both Fama-MacBeth regressions and pooled time-series regressions. The argumentation for running two different regressions, as in Pedersen *et al.* (2021), is that the standard errors for the two methods are estimated differently. The following section further explains the two regression models.

5.2 Fama-MacBeth Regressions

The first method is the two-stage Fama-MacBeth regression to estimate the risk premia of ESG scores. This method successfully works with panel data and multiple assets'

(Fama and MacBeth, 1973). In the first stage, we regress n assets returns against m suggested risk factors to determine each asset's exposure to the risk factors shown in the regressors beta. This stage is done using time-series regressions as in Equation 11

$$R_{i,t} = \alpha_i + \boldsymbol{\beta}_i \boldsymbol{F}_t + \epsilon_{i,t}. \tag{11}$$

Where $R_{i,t}$ is the calculated return for asset i at time t. $\mathbf{F_t}$ is the factors at time t and $\boldsymbol{\beta_i}$ its factor exposure. We then regress all asset returns against the same factors' F. In the second stage, we run cross-sectional regressions on the assets returns for T periods against our estimated betas from Equation 11. This method will help us determine the risk premium for our selected factors.

$$R_{i,t} = \hat{\boldsymbol{\beta}} \lambda_t + \alpha_t \tag{12}$$

We still have $R_{i,t}$ on the left-hand side, but we use the estimated betas as regressors, giving us our λ coefficients for each factor. We then calculate our estimated risk premiums using the coefficients by averaging the coefficients as in Fama and MacBeth (1973),

$$\hat{\boldsymbol{\lambda}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\lambda}_t. \tag{13}$$

$\hat{\boldsymbol{\lambda}}_m$ will be a vector of the estimated risk premiums of our m factors. Following Green *et al.* (2017) in the usage of Fama-MacBeth regressions to generate the most potent independent identifiers of average return. The reasoning for running Fama-MacBeth regressions instead of construction of portfolios is similar to Halbritter and Dorfleitner (2015) that doing so includes the whole data set without portfolio sorting, as only including firms with high or low ESG scores would leave much explanatory data behind.

5.3 Time Series Regressions

Besides Fama-MacBeth regressions, we also apply time-series regressions, similar to Pedersen *et al.* (2021). Since we are working with cross-sectional data over multiple years, we control for unobserved variables, thus decreasing our bias. The performed regressions use the R package $\mathbb{PLM}$, where we set the dates and firms as fixed. One can denote the general panel data regression with return on the left-hand side as,

$$R_{i,t} = \alpha_i + \boldsymbol{X'}_{i,t}\boldsymbol{\beta} + \epsilon_{i,t}. \tag{14}$$

Where returns for firm i across different time periods equals: $R'_{i,t} = R_{i.t}, ..., R_{i.T}$ and $\boldsymbol{X}_{i,t}$ is a vector of explanatory characteristics for firm i at period t.

5.4 Robustness Tests

Our panel data consists of relatively few years and various companies. In these circumstances, we intend to assume we are violating the assumptions of homoskedasticity and no serial correlation. For instance, if a company has had a high ESG score in previous years, it may also have a high ESG score this year. Furthermore, the variance across elements is rarely identical, and heteroscedasticity occurs in almost all types of cross-sectional data. The standard and most efficient way to correct these problems are using Newey West standard errors in all time series regressions (Gujarati, 2011).

However, we still perform some tests, and Table 6 in the Appendix shows the anticipated results using the Breusch-Pagan test for heteroskedasticity, where the null hypothesis claims homoskedasticity, which we reject. The Breusch-Godfrey test is for serial correlation, where the null hypothesis is that there is no serial correlation. The Breusch-Godfrey test can be shown in Table 7 in the Appendix, and we fail to reject the null hypothesis. From our Correlation Matrix in Table 8 in the Appendix, the financial variables suffer low correlation. The ESG variables have higher correlations, but there will be no problem in our regression analysis since we include them separately.

6 Results

Table 2 presents our main results from our pooled time-series regressions and Fama-MacBeth regressions on yearly excess returns with ESG scores and return predictors. In Regression (1) and Regression (2), we see the ESG score and return predictors on yearly excess return without the Sin stock dummy variable. In contrast, we include the Sin stock dummy variable in Regression (3) and Regression (4). In total, we have 4086 observations over our estimated time period.

Our findings are that the ESG score significantly explains yearly excess return, similar to Halbritter and Dorfleitner (2015) and Maiti (2021). This relationship holds in all four regressions and is negative at a 0.1% significance level. Thus, firms with a high ESG score perform lower yearly excess return *ceteris paribus*. This result is in line with Pedersen *et al.* (2021) for ESG-motivated investors and Hypothesis 1.

The pooled time-series regressions estimate a more negative coefficient (-0.0037) for ESG than the Fama-MacBeth regressions (-0.0025). That is, an increase in the ESG score by one is associated with a decrease in yearly excess return by 37 basis points according to our pooled time-series estimate and by 25 basis points according to our Fama-MacBeth estimate, *ceteris paribus*. Other return predictors such as *LSIZE* and *B/M* shows significance in all regressions in Table 2. The pooled time-series regressions show that *BETA* and *INV* also is significant.

In Regression (3) and Regression (4), we find no evidence of the *Sin Stock premium*; hence, the *Sin Stock anomaly* cannot explain the negative effect of the ESG risk score. Additionally, the Sin stock dummy variable does not contribute to higher explanatory power for our in-sample estimates. We see little or no effect on the coefficients when controlling for the *Sin Stock anomaly*. This result is not in line with Hong and Kacperczyk (2009) and Pedersen *et al.* (2021), who found a positive and significant effect on excess return for Sin stocks. By the results in Regression (3) and Regression (4), we cannot assume Hypothesis 2 to be true. Thus, in our tests, the *Sin Stock anomaly* does not exist.

Table 2: Regressions of yearly excess return on ESG score

	(1)	(2)	(3)	(4)
BETA	0.0412**	0.0166	0.0412**	0.0164
	(0.0138)	(0.0537)	(0.0138)	(0.0534)
LSIZE	0.0266***	0.0260***	0.0266***	0.0259***
	(0.0065)	(0.0062)	(0.0065)	(0.0060)
B/M	−0.1256***	−0.2074***	−0.1254***	−0.2072***
	(0.0240)	(0.0352)	(0.0240)	(0.0342)
INV	0.6587***	0.2462	0.6608***	0.2415
	(0.1705)	(0.3125)	(0.01708)	(0.3116)
PROF	0.0141	−0.1590	0.0139	−0.1580
	(0.0806)	(0.1311)	(0.0806)	(0.1296)
ESG	−0.0037***	−0.0025***	−0.0037***	−0.0025***
	(0.0003)	(0.0004)	(0.0003)	(0.0004)
SIN			0.0057	−0.0004
			(0.0246)	(0.0295)
R^2	0.061	0.104	0.061	0.104
Num. obs.	4086	4086	4086	4086
Estimation method	Pooled	FM	Pooled	FM

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The dependent variable represents yearly excess returns, and the independent variables are in the left column. Regression (1) and (2) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 ESG_{i,t} + \epsilon_{i,t}$ and Regression (3) and (4) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 ESG_{i,t} + \theta_2 SIN_{i,t} + \epsilon_{i,t}$. Where $\boldsymbol{X}_{i,t}$ is a vector that includes the firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and *INV*; $ESG_{i,t}$ is the ESG score for firm i in period t; $SIN_{i,t}$ is a dummy variable that equals one if the stock is a Sin stock and zero otherwise; $\epsilon_{i,t}$ is the measurement error. Pooled stands for pooled time-series regressions and FM stands for Fama-MacBeth regressions. Significance levels are as $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Robust standard errors in parentheses.

The pooled time-series and Fama-MacBeth regressions for E, S, and G separately are shown in Table 3 below and establish a negative relationship between all pillar scores and yearly excess return.

Regression (5) and Regression (6) include the Environmental pillar score variable E and all firm characteristics as control variables and indicate that the negative relationship between yearly excess return and Environmental pillar scores is true at a 0.1% significance level using both methodologies. Namely, an increase in the Environmental pillar score by one is associated with a decrease in yearly excess return by 24 basis points for the pooled time-series estimate and by 17 basis points for the Fama-

MacBeth estimate, *ceteris paribus.* Thus, holding environmental-friendly assets on the S&P 500 Index is associated with lower yearly excess returns and demonstrates the theoretical framework of ESG-motivated investors accepting lower returns of holding certain assets, *ceteris paribus.* At the same time, Pedersen *et al.* (2021) used a carbon emission proxy for the Environmental score. They found weak or little evidence of an *Environmental risk premium*, similar to Bolton and Kacperczyk (2021), who found a carbon premium that disappeared after controlling for industry composition. Hence, this result is in line with Bolton and Kacperczyk (2021) and Pedersen *et al.* (2021) that an environmental premium exists; besides, we test for the whole pillar and confirm the premium at a more vital significance level.

Regression (7) and Regression (8) in Table 3 denote the pooled time-series regression and Fama-MacBeth regression for the Social pillar score variable S and all firm characteristics as control variables. Consistent with previous research (Hong and Kacperczyk, 2009, Pedersen *et al.*, 2021), a negative relationship was found between the Social pillar score and yearly excess return in Regression (7) and Regression (8), respectively, where the Fama-MacBeth regression gives a less negative estimate. That is, an increase in the Social pillar score by one is associated with a decrease in yearly excess return by 30 basis points for the pooled time-series estimate and by 19 basis points for the Fama-MacBeth estimate. Our results indicate that a *Social risk premium* exists. We can reject the null hypothesis at a 0.1% significance level and assume the negative relationship to be true, *ceteris paribus.* Even though the methodology by Pedersen *et al.* (2021) is different, as they sorted stocks depending on Sin stock or not and created two groups, the results match our prediction.

The last regressions, Regression (9) and Regression (10), in Table 3 include the Governance pillar score variable G and all firm characteristics as control variables and predict significant negative results. The pooled time-series regression for G, Regression (9), estimates the Governance impact to be negative (-0.0018) and significant at the 0.1% level, while the Fama-Macbeth regression for G, Regression (10), shows a weak negative coefficient (-0.0010) only significant on the 5% level. More precisely, an increase in the Governance pillar score by one is associated with a decrease in yearly excess return by 18 basis points for the pooled time-series estimate and 10 basis points for the Fama-MacBeth estimate, *ceteris paribus*. Remember Harvey *et al.* (2016), who argue for t-statistics larger than 3.0; in Regression (10), we should reject the 2.0 t-statistics of the Governance pillar score as a return predictor. The prediction is not in line with Bebchuk *et al.* (2013) findings of no evidence of a relationship between abnormal returns and governance indices nor the earlier positive relationship. Providing different results can be due to other proxies for governance; they test the effect close to good-governance announcements and possess a different period while we, same as with E and S, focus on the pillar score from Refinitiv Eikon.

Nevertheless, with ESG scores successfully explaining the yearly excess return, we stand with the predecessors Friede *et al.* (2015), Hübel and Scholz (2020), and Maiti (2021) regarding studies that show ESG scores function as return predictors. For all our tests in Table 2 and Table 3, this holds, with only Regression (10) in Table 3 showing slightly weaker significance for the Governance pillar score (5%).

Table 3: Regressions of yearly excess return on ESG pillar scores

	(5)	(6)	(7)	(8)	(9)	(10)
BETA	0.0456^{***}	0.0209	0.0444^{***}	0.0182	0.0434^{**}	0.0179
	(0.0137)	(0.0543)	(0.0135)	(0.0540)	(0.0138)	(0.0540)
LSIZE	0.0267^{***}	0.0292^{***}	0.0221^{***}	0.0229^{***}	0.0007	0.0099^{*}
	(0.0065)	(0.0058)	(0.0069)	(0.0061)	(0.0055)	(0.0046)
B/M	-0.1149^{***}	-0.1987^{***}	-0.1437^{***}	-0.2205^{***}	-0.1325^{***}	-0.2171^{***}
	(0.0250)	(0.0357)	(0.0242)	(0.0368)	(0.0240)	(0.0336)
INV	0.7558^{***}	0.2922	0.6197^{***}	0.2196	0.7350^{***}	0.2840
	(0.1732)	(0.3050)	(0.1705)	(0.3108)	(0.1676)	(0.3101)
PROF	−0.0067	−0.1770	0.0102	−0.1631	0.0076	−0.1832
	(0.0820)	(0.1340)	(0.0814)	(0.1351)	(0.0809)	(0.1321)
E	-0.0024^{***}	-0.0017^{***}				
	(0.0002)	(0.0002)				
S			-0.0030^{***}	-0.0019^{***}		
			(0.0003)	(0.0003)		
G					-0.0018^{***}	-0.0010^{*}
					(0.0003)	(0.0005)
R^2	0.057	0.104	0.054	0.101	0.037	0.101
Num. obs.	4086	4086	4086	4086	4086	4086
Estimation method	Pooled	FM	Pooled	FM	Pooled	FM

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The dependent variable represents yearly excess returns, and the independent variables are in the left column. Regression (5) and (6) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 E_{i,t} + \epsilon_{i,t}$, Regression (7) and (8) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 S_{i,t} + \epsilon_{i,t}$ and Regression (9) and (10) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 G_{i,t} + \epsilon_{i,t}$. Where $\boldsymbol{X}_{i,t}$ is a vector that includes the earlier described firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and *INV*; $E_{i,t}$, $S_{i,t}$, and $G_{i,t}$ is the separately conducted pillar scores from Refinitiv Eikon; and $\epsilon_{i,t}$ is the measurement error. Pooled stands for pooled time-series regressions and FM stands for Fama-MacBeth regressions. Significance levels are shows as $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Robust standard errors in parentheses.

6.1 Robustness

As in Bolton and Kacperczyk (2021), we then include the General Industry Classification System to explore how the sector fixed effects change our estimates as a robustness test. We create sector dummy variables and continue with pooled time-series regressions with firm and date as fixed variables. We cluster our robust standard errors on a firm level. The classification and distribution of companies are in Table 9 in the Appendix.

This conservative method is as in Hong and Kacperczyk (2009) that the errors are conditional on independent variables correlated within sectors. The results in Table 4 show little or no impact on our ESG risk score factors compared to earlier tests. We still have significant results at 0.1%; the coefficients are slightly smaller for ESG, E, and S and, on the hand, bigger for G. We then have reasons to believe that our results are robust when controlling for firm sectors.

In a second robustness test, we split the sample into two categories of firms, those with higher ESG scores than the median and those with ESG scores under the median. This split is to generate a comparison between "greener" assets and "browner" assets. The idea is that all explanation exists within the "brown" assets; hence assets above the median may show no relationship. We run pooled time-series regressions with firm and date as fixed. We are presenting the results in Table 10 in the Appendix, and one can see that our results are unaffected by this split of the sample since both significance levels and coefficients remain robust.

Table 4: Fixed effects regressions of yearly excess return on ESG scores

	(11)	(12)	(13)	(14)
BETA	0.0385^{**}	0.0448^{***}	0.0384^{**}	0.0388^{**}
	(0.0130)	(0.0130)	(0.0130)	(0.0131)
LSIZE	0.0308^{***}	0.0304^{***}	0.0268^{***}	0.0042
	(0.0054)	(0.0056)	(0.0053)	(0.0049)
B/M	-0.1018^{***}	-0.0911^{***}	-0.1121^{***}	-0.1157^{***}
	(0.0223)	(0.0225)	(0.0223)	(0.0226)
INV	1.2774^{***}	1.2936^{***}	1.3238^{***}	1.3479^{***}
	(0.1730)	(0.1736)	(0.1732)	(0.1751)
PROF	−0.1006	−0.1222	−0.1002	−0.0690
	(0.0830)	(0.0834)	(0.0832)	(0.0841)
ESG	-0.0038^{***}			
	(0.0003)			
E		-0.0023^{***}		
		(0.0002)		
S			-0.0031^{***}	
			(0.0003)	
G				-0.0017^{***}
				(0.0002)
R^2	0.079	0.071	0.075	0.054
Industry fixed effect	Yes	Yes	Yes	Yes
Num. obs.	4086	4086	4086	4086
Etimation method	Pooled	Pooled	Pooled	Pooled

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The dependent variable represents yearly excess returns, and the independent variables are in the left column. Regression (11) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 ESG_{i,t} + \theta_2 \boldsymbol{Industry}_{i,t} + \epsilon_{i,t}$, Regression (12) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 E_{i,t} + \theta_2 \boldsymbol{Industry}_{i,t} + \epsilon_{i,t}$, Regression (13) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 S_{i,t} + \theta_2 \boldsymbol{Industry}_{i,t} + \epsilon_{i,t}$ and Regression (14) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 G_{i,t} + \theta_2 \boldsymbol{Industry}_{i,t} + \epsilon_{i,t}$. Where $\boldsymbol{X}_{i,t}$ is a vector that includes the earlier described firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and *INV*; $ESG_{i,t}$ is the ESG score for firm i in period t; $E_{i,t}$, $S_{i,t}$, and $G_{i,t}$ is the separately conducted pillar scores from Refinitiv Eikon; $\boldsymbol{Industry}_{i,t}$ is a vector with dummy variables for each industry that equals one if the stock is included in the specific industry and zero otherwise; $\epsilon_{i,t}$ is the measurement error. Pooled stands for pooled time-series regressions. Significance levels are as $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Robust standard errors clustered on firm levels in parentheses.

7 Conclusion

We imply that our results are two-fold; first, we find supporting evidence to the theoretical framework of ESG-motivated investors; second, ESG scores from Refinitiv Eikon explain yearly excess return where both our pooled time-series regressions and the Fama-MacBeth regressions show a significant negative relationship. Hence, we show that on the S&P 500 Index, investors accept lower returns for holding a higher ESG portfolio. This result is in line with the theory of Pástor *et al.* (2021) and Pedersen *et al.* (2021).

We prove that the highest possible expected excess return is not the only factor for investors when making investment decisions since they gain utility for their tastes for assets. Our attempt to retain all possible information using a cross-sectional approach successfully results in a significant negative relationship between yearly excess return and ESG scores. We contribute by increasing the explanatory power when including the ESG score as a return predictor in our regression models.

With additional tests, we show that the pillar scores for Environmental and Social are in line with the theory of Hong and Kacperczyk (2009) and Bolton and Kacperczyk (2021). These results indicate that the Environmental and Social pillar scores successfully work as proxies for carbon emissions and unethical businesses, strengthening our hypothesis of an *Environmental risk premium* and a *Social risk premium.* Investors are then willing to trade off excess return to decrease their portfolios' regulatory risk since these businesses are more exposed to new policies.

However, when we control for the *Sin Stock anomaly*, we fail to show significant effects for Sin stocks; we argue that this may be to the lack of observations of firms in the industries specified by Fama and French (1997). On the other hand, the Governance pillar, which shows a significant negative relationship, is incorrect compared to Bebchuk *et al.* (2013). We argue that this could mean that the Governance pillar score does not work as a proxy for corporate governance and that investors today incorporate good corporate governance as something obvious in their investment decisions.

7.1 Limitations

As Halbritter and Dorfleitner (2015) stated, the choice of ESG score provider may affect the results since the rating providers estimate firms' emissions, green initiatives, and social contributions differently. The risk is that larger firms, as in our sample, will have better scores since they can work on their CSR issues. The problem with the ESG scores and the relatively short time period of available data for many firms can be troublesome. Our limitation is to yearly observations since we need variation in the data, and the ESG scores are updated yearly. The effect may have been different with the possibility of quarterly or monthly scores from other providers.

Regarding the *Sin Stock anomaly* where we find no effect on yearly excess return, one needs to recall that, unfortunately, only 18 firms were available in our sample. While we follow the classification used by Fama and French (1997), one could also include other industries such as *Weapon & Defence* and *Oil & Gas*, which may give different results. We also limit our sample to listed shares on the S&P 500 Index, which includes the biggest listed companies in the world. One may find other results with a broader sample of firms of various sizes and from different geographical regions.

7.2 Further Research

We conduct a cross-sectional approach to determine the ESG risk score factor in the US. Our results intend that ESG investing is due to investor preferences and state a negative relationship between ESG scores and yearly excess return on the S&P 500 Index. We suggest further research within this field, and one way is to compare the ESG score between different markets and rating providers. By comparing the ESG risk score factor between stock indices, for example, the S&P 500 Index, FTSE 100 Index, DAX Index, Nikkei, and Hang Seng. Interesting will be if investors' preferences differ between these international stock indices or if using other rating providers affects the results.

8 References

Banz, R. W. (1981), 'The relationship between return and market value of common stocks', *Journal of Financial Economics* **9**(1), 3–18.

Bebchuk, L. A., Cohen, A. and Wang, C. C. (2013), 'Learning and the disappearing association between governance and returns', *Journal of Financial Economics* **108**(2), 323–348.

Bolton, P. and Kacperczyk, M. (2021), 'Do investors care about carbon risk?', *Journal of Financial Economics* **142**(2), 517–549.

Carhart, M. (1997), 'On persistence in mutual fund performance', *Journal of Finance* **52**(1), 57–82.

Daniel, K. and Titman, S. (1997), 'Evidence on the characteristics of cross sectional variation in stock returns', *The Journal of Finance* **52**(1), 1–33.

DeMarzo, P. and Berk, J. (2016), *Corporate Finance*, Pearson Education.

Fama, E. F. and French, K. R. (1997), 'Industry costs of equity', *Journal of Financial Economics* **43**(2), 153–193.

Fama, E. F. and French, K. R. (2007), 'Disagreement, tastes, and asset prices', *Journal of Financial Economics* **83**(3), 667–689.

Fama, E. F. and MacBeth, J. D. (1973), 'Risk, return, and equilibrium: Empirical tests', *The Journal of Political Economy* **81**, 607–636.

Fama, E. and French, K. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1), 3–56.

Fama, E. and French, K. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1–22.

Friede, G., Busch, T. and Bassen, A. (2015), 'Esg and financial performance: aggregated evidence from more than 2000 empirical studies', *Journal of Sustainable Finance & Investment* **5**(4), 210–233.

Galema, R., Plantinga, A. and Scholtens, B. (2008), 'The stocks at stake: Return and risk in socially responsible investment', *Journal of Banking & Finance* **32**(12), 2646–2654.

Green, J., Hand, J. and Zhang, F. (2017), 'The characteristics that provide independent information about average u.s. monthly stock returns', *The Review of Financial Studies* **30**(12), 4389–4436.

Gujarati, D. N. (2011), *Econometrics by example*, Vol. 1, Palgrave Macmillan New York.

Halbritter, G. and Dorfleitner, G. (2015), 'The wages of social responsibility—where are they? a critical review of esg investing', *Review of Financial Economics* **26**, 25–35.

Harvey, C. R., Liu, Y. and Zhu, H. (2016), '. . . and the cross-section of expected returns', *The Review of Financial Studies* **29**(1), 5–68.

Hong, H. and Kacperczyk, M. (2009), 'The price of sin: The effects of social norms on markets', *Journal of Financial Economics* **93**(1), 15–36.

Hübel, B. and Scholz, H. (2020), 'Integrating sustainability risks in asset management: The role of esg exposures and esg ratings', *Journal of Asset Management* **21**(1), 52–69.

Jegadeesh, N. and Titman, S. (1993), 'Returns to buying and selling losers: Implications for stock market efficiency', *Journal of Finance* **48**(1), 65–91.

Kahneman, D. and Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**(2), 263–291.

Litner, J. (1965), 'Security prices, risk, and maximal gains from diversification', *Journal of Finance* **20**(4), 587–615.

Maiti, M. (2021), 'Is esg the succeeding risk factor?', *Journal of Sustainable Finance & Investment* **11**(3), 199–213.

Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.

Mossin, J. (1966), 'Equilibrium in a capital asset market', *Econometrica: The Econometric Society* **34**(4), 768–783.

Naffa, H. and Fain, M. (2022), 'A factor approach to the performance of esg leaders and laggards', *Finance Research Letters* **44**, 102073.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2021), 'Sustainable investing in equilibrium', *Journal of Financial Economics* **142**(2), 550–571.

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

Refinitiv (2021), 'Environmental, soical and governace scores from refintiv'.
URL: *https://www.refinitiv.com/content/dam/marketing/en_us/documents /methodology/refinitiv-esg-scores-methodology.pdf*

Revelli, C. and Viviani, J.-L. (2015), 'Financial performance of socially responsible investing (sri): what have we learned? a meta-analysis', *Business Ethics: A European Review* **24**(2), 158–185.

Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *Journal of Financial Economics* **9**(3), 424–442.

9 Appendix

Table 5: Categories for ESG scores

Category	Category weight
Emissions	0.15
Resource use	0.15
Innovation	0.13
Community	0.09
Human rights	0.05
Product responsibility	0.04
Workforce	0.13
Shareholders	0.05
CSR strategy	0.03
Management	0.17

Note: Information from Refinitiv Eikon, including the different categories and weights for their ESG scores.

Table 6: Breusch-Pagan Test for Heterskedasticity

Breusch-Pagan Test	ESG	E	S	G
Chi2-value	72	69	72	73
P-value	0.00***	0.00***	0.00***	0.00***
Degrees of freedom	6	6	6	6

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The dependent variable represents yearly excess return, and Regression (ESG) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 ESG_{i,t} + \epsilon_{i,t}$, Regression (E) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 E_{i,t} + \epsilon_{i,t}$, Regression (S) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 S_{i,t} + \epsilon_{i,t}$ and Regression (G) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 G_{i,t} + \epsilon_{i,t}$. Where $\boldsymbol{X}_{i,t}$ is a vector that includes the earlier described firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and *INV*; $ESG_{i,t}$ is the ESG score for firm i in period t; $E_{i,t}$, $S_{i,t}$, and $G_{i,t}$ is the separately conducted pillar scores from Refinitiv Eikon; and $\epsilon_{i,t}$ is the measurement errors. The test rejects the null hypothesis and assume variance among residuals in all regressions. Significance levels are shown as $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 7: Breusch-Godfrey/Wooldridge Test for Serial Correlation

Breusch-Godfrey Test	ESG	E	S	G
LM-test	0.7	0.3	0.6	0.7
P-value	0.4	0.6	0.4	0.4
Degrees of freedom	1	1	1	1

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The dependent variable represents yearly excess return, and Regression (ESG) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 ESG_{i,t} + \epsilon_{i,t}$, Regression (E) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 E_{i,t} + \epsilon_{i,t}$, Regression (S) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 S_{i,t} + \epsilon_{i,t}$ and Regression (G) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma X}_{i,t} + \theta_1 G_{i,t} + \epsilon_{i,t}$. Where $\boldsymbol{X}_{i,t}$ is a vector that includes the earlier described firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and *INV*; $ESG_{i,t}$ is the ESG score for firm i in period t; $E_{i,t}$, $S_{i,t}$, and $G_{i,t}$ is the separately conducted pillar scores from Refinitiv Eikon; and $\epsilon_{i,t}$ is the measurement errors. The test fails to reject the null hypothesis and assume that no serial correlation can be found in the regressions. Significance levels are shown as $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 8: Correlation Matrix

	BETA	LSIZE	B/M	INV	PROF	ESG	E	S	G
BETA	1.000								
LSIZE	-0.063	1.000							
B/M	0.099	0.013	1.000						
INV	-0.117	0.005	0.000	1.000					
PROF	-0.024	0.041	-0.466	-0.024	1.000				
ESG	-0.071	0.497	0.054	-0.023	0.003	1.000			
E	-0.042	0.518	0.127	0.022	-0.046	0.859	1.000		
S	-0.053	0.487	-0.008	-0.052	0.032	0.883	0.735	1.000	
G	-0.067	0.204	0.088	0.005	-0.019	0.677	0.396	0.367	1.000

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The financial variables suffer low correlation with each other. The ESG and E, S, and G separately tend to have a higher correlation; these variables will not be in the same regressions; hence no multicollinearity problem can be found.

Table 9: Firms by Industry Sector

GIC	Sector	# of firms
10	Energy	20
15	Materials	26
20	Industries	68
25	Consumer Discretionary	57
30	Consumer Staples	29
35	Health Care	61
40	Financials	61
45	Information Technology	69
50	Communication Services	23
55	Utilities	28
60	Real Estate	28

Note: Contains information about the number of firms belonging to each sector. Information Technology and Industries are the two largest sectors containing 69 and 68 firms, respectively. Energy is the smallest sector in the data set and contains 20 firms.

Table 10: Regressions of yearly excess return on ESG scores, and pillar scores for "green" and "brown" assets

	Green				Brown			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
BETA	0.0509***	0.0521***	0.0548***	0.0570***	0.0325	0.0405*	0.0322	0.0280
	(0.0142)	(0.0143)	(0.0142)	(0.0144)	(0.0179)	(0.0179)	(0.0180)	(0.0180)
LSIZE	0.0305***	0.0278***	0.0306***	0.0089	0.0178	0.0225*	0.0098	−0.0059
	(0.0062)	(0.0063)	(0.0062)	(0.0058)	(0.0092)	(0.0095)	(0.0092)	(0.0087)
B/M	−0.1092***	−0.0939***	−0.1146***	−0.1212***	−0.1402***	−0.1345***	−0.1738***	−0.1485***
	(0.0267)	(0.0271)	(0.0267)	(0.0271)	(0.0310)	(0.0311)	(0.0308)	(0.0315)
INV	0.8030***	0.9105***	0.7688***	0.9206***	0.5174*	0.6229**	0.4946*	0.5911**
	(0.1919)	(0.1927)	(0.1922)	(0.1947)	(0.2196)	(0.2194)	(0.2215)	(0.2213)
PROF	0.0802	0.0685	0.1228	0.0506	−0.0511	−0.0760	−0.0813	−0.0284
	(0.1021)	(0.1028)	(0.1023)	(0.1039)	(0.1154)	(0.1154)	(0.1162)	(0.1166)
ESG	−0.0041***				−0.0040***			
	(0.0004)				(0.0005)			
E		−0.0023***				−0.0026***		
		(0.0003)				(0.0003)		
S			−0.0034***				−0.0025***	
			(0.0004)				(0.0005)	
G				−0.0014***				−0.0018***
				(0.0003)				(0.0004)
R^2	0.069	0.057	0.069	0.037	0.052	0.052	0.040	0.035
Num. obs.	2148	2148	2148	2148	1937	1937	1937	1937
Estimation method	Pooled	Pooled	Pooled	Pooled	Pooled	Pooled	Pooled	Pooled

Note: Contains data from S&P 500 Index between 2010 and 2020, including 470 constituents. The dependent variable represents yearly excess returns, and the independent variables are in the left column. Regression (1) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 ESG_{i,t} + \epsilon_{i,t}$ and Regression (2) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 E_{i,t} + \epsilon_{i,t}$ and Regression (3) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 S_{i,t} + \epsilon_{i,t}$ and Regression (4) are $R_{i,t} - R_t^f = \alpha_i + \boldsymbol{\gamma} \boldsymbol{X}_{i,t} + \theta_1 G_{i,t} + \epsilon_{i,t}$. Where $\boldsymbol{X}_{i,t}$ is a vector that includes the earlier described firm characteristics *BETA*, *LSIZE*, *B/M*, *PROF*, and *INV*; $ESG_{i,t}$ is the ESG score for firm i in period t; $E_{i,t}$, $S_{i,t}$, and $G_{i,t}$ is the separately conducted pillar scores from Refinitiv Eikon; and $\epsilon_{i,t}$ is the measurement errors. Pooled stands for pooled time-series regressions. Significance levels are shows as $^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Robust standard errors in parentheses.