Sustainability Risk as an Asset Pricing Factor[1]

Maria Stella Kusuma Dewi

m.mariastellakusumadewi@students.uu.nl

Student Number: 7281536

Thesis Supervisor: Dr. Yunhui Han

Co-reader: Dr. Kwabena Aboah Addo

Abstract

This thesis examines whether widely used asset pricing models can be improved by adding a sustainability risk factor. It constructs and assesses long-short sustainability factors in the style of Small Minus Big (SMB) and High Minus Low (HML) as in Fama & French (1993). The sustainability factors are based on the Total ESG Risk and the three individual ESG pillars' risk for firms in the S&P500. The Total ESG Risk factor, combined with the Market, SMB, and HML factors, results in more accurate descriptions of stock returns than the CAPM and Fama-French Three Factor model. The resulting implication for finance practitioners is that omission of sustainability risk in their asset pricing models might lead to misestimations of required rates of returns.

JEL: G12, Q51

Keywords: Asset Pricing; Sustainability; Factors; ESG

---

[1] I want to thank my supervisor Dr. Yunhui Han, for her helpful feedback and guidance, especially during the proposal period. I am very grateful to my family for their full encouragement and constant support. Special thank you to Sian Tee for his patience and encouragement. And to everyone who has helped me along the way, thank you for making this achievement possible.

Table of Contents

1 Introduction

This thesis examines whether widely used asset pricing models can be improved by adding a sustainability risk factor. Asset pricing models are applied across the financial industry by practitioners to determine what rate of return can be expected, given the level of risks taken. Sustainability risks are not yet widely taken into account for this purpose.

The Capital Asset Pricing Model (CAPM) – developed by Treynor (1961, 1962), Sharpe (1964), Lintner (1965a,b), and Mossin (1966) – is still one of the most influential asset pricing models, around six decades after its introduction. Asset pricing models are judged based on their ability to describe asset returns without leaving a significant part of the returns unexplained. Despite its popularity, the CAPM is not the best in this regard. For example, Fama & French (1993) show that their Fama-French Three Factor model (FF3) results in lower pricing errors. These models and the numerous that followed them did not yet lead to consensus on a single asset pricing model that performs well under varying circumstances: a sign that essential factors to complete the puzzle are still missing.

A part of the missing pieces can potentially be found in the recently rapidly evolving research on sustainability risk. Sustainability – commonly divided into the three pillars Environment, Social, and Governance (ESG) – is a theme that is shifting to the centre of attention society-wide, including within finance. Sustainability risks can – through e.g., physical, reputational, or regulatory consequences – affect a company's (expected) ability to stay profitable. This can in turn affect the company's stock price (Hong & Kacperczyk., 2009; Liou., 2018). However, sustainability as a factor in an asset pricing context is still under-examined. This thesis contributes to addressing the gaps in asset pricing and sustainability risk research, by examining the following research question:

To what extent can sustainability risk factors augment the FF3 model to explain stock returns more accurately than benchmarks CAPM and FF3?

To answer this question, I construct four sustainability factors, i.e., one based on Total ESG Risk and three based on each individual pillar Environmental, Social, and Governance Risk. For constructing and assessing these factors, I follow a similar approach to Fama & French (1993). Each of the four sustainability factors are separately added to the FF3 model, resulting in four different augmented versions of the FF3.

Along three hypotheses, the factors are sequentially assessed for (i) their individual explanatory power of returns, (ii) their contribution to a model that can explain more return variation than benchmarks, and (iii) their ability to fully reduce the pricing errors of FF3 to zero. The results indicate that adding a factor based on Total ESG Risk leads to more accurate description of stock returns, while the other three sustainability factors do not seem to improve the FF3 model.

Besides addressing the academic gap, this research also has social relevance, given the widespread use of asset pricing models in the financial industry. An improved asset pricing model enables better-informed decision making.

For example, asset pricing models can be used for assessing a firm's weighted average cost of capital (WACC) – an important input to the valuation of the firm – and hence can affect mergers and acquisitions (M&A). Furthermore, the performance of investment managers (e.g., for pension funds) is often assessed based on the return they achieved above and beyond a benchmark. By using a valid asset pricing model, this benchmark can be based on the risks taken by the manager. A benchmark that does not consider all relevant risks can result in incentivising investment managers to take risks that are not justified by their return. By suggesting potential improvements to the widely used FF3 model, this thesis can hence contribute to improving financial practices.

To construct and test the sustainability factor models, this thesis first reviews and formalises the theoretical framework it is built upon, which lead to the three hypotheses in chapter 2. To test these hypotheses empirically, chapter 3 describes how the required data is collected and transformed into the factor returns. Chapter 4 describes both the tests and their corresponding results. Finally, chapter 5 places the findings in the larger academic and social context, provides several limitations, and poses suggestions for future research.

2 Theory and Hypotheses

2.1 Asset Pricing Factor Models

Asset pricing is the academic field involved with the examination of price movements of financial assets. Factor models have been playing a dominant role within asset pricing research over the last decades. The fundamental concept underlying these models is that there is a limited set of variables – factors – that together can fully explain the expected returns of every asset. These factors represent systematic (i.e., non-diversifiable) sources of risk in the economy, for which investors can expect a return when exposing themselves to them. That is, the factors are common for the full market, but each asset can have a unique exposure to the factors.

The most common factor models in academic literature assume expected returns to be a linear combination of expected factor returns and asset-specific factor exposures. In mathematical terms, these general K-factor models are defined as

$$\mathrm{E}(\mathrm{r_i}) = \beta_{1;i}E(f_1) + \cdots + \beta_{K;i}E(f_K), \tag{1}$$

where $\mathrm{E}(\mathrm{r_i})$ indicates the expected return of an asset i in excess of the risk-free rate of return; $E(f_1), \ldots, E(f_K)$ indicate the expected returns of the K factors; and $\beta_{1;i}, \ldots, \beta_{1;K}$ indicate the exposures of asset i to each of the respective K factors. This model can be applied to any return periodicity, e.g., daily, monthly, or yearly returns. However, all returns $\mathrm{r_i}, f_1, \ldots, f_K$ within a model specification must have the same periodicity.

An essential aspect of the set of factor models described by (1) is that it does not contain an intercept. Many empirical tests of asset pricing models therefore amount to assessing whether estimated intercepts – referred to as pricing errors – are indeed statistically indistinguishable from zero, after fitting the model to observed return data. With the aim of obtaining an empirically valid factor model with zero pricing errors, numerous variants have been proposed in academic literature. The following subsections review some of the most notable of them.

2.1.1 Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is a single factor model, with the excess return on the overall market being its only factor. The CAPM is developed by Treynor (1961, 1962), Sharpe (1964), Lintner (1965a,b), and Mossin (1966). The model is built upon the foundations laid in the modern portfolio theory by Markowitz (1952) that describes mean-variance efficient portfolio construction.

Equation (1) under the CAPM becomes

$$E(r_i) = \beta_{M;i} E(r_M), \tag{2}$$

where $E(r_M)$ indicates the expected value-weighted return of the full market in excess of the risk-free rate; $\beta_{M;i}$ indicates the exposure of an asset i to the market factor; and $E(r_i)$ again indicates the expected excess return of asset i.

Since its introduction in the 1960's until this day, the CAPM has been one of the most widely used factor models, but it is not free from critique, both theoretical and empirical. On the one hand, the model's theoretical assumptions have been criticised for being unrealistic. On the other hand, empirical research in numerous markets and time periods have indicated that the CAPM does not explain returns without pricing errors.

Some of the theoretical critique challenges the CAPM's assumption that all investors are rational with homogeneous expectations of risk and return. The field of behavioural finance examines some of the biases present in investors' expectations and aims to correct for these biases, for example, in the overconfidence-adjusted asset pricing models by Daniel, Hirshleifer, and Subrahmanyam (2001).

The empirical evidence against the CAPM's validity finds examples of anomalies, i.e., observations that are inconsistent with the model's predictions. Two of the most prominent examples are the size and value anomalies, which result in the Fama-French Three Factor Model, as explained in the next subsection.

Regardless of the critiques on the model, the CAPM is still widely used by finance practitioners due to its simple nature. For example, for evaluating an investment manager's performance or when calculating a company's required rate of return for valuation purposes. Due to its wide use and influence in practice, the CAPM also remains an important benchmark model in academics, including in this thesis.

2.1.2 Fama-French Three Factor Model

The Fama-French Three Factor Model (FF3) extends the CAPM with two additional factors, related to size and value. It is based on the empirical anomalies that the stock market tends to be outperformed by (i) stocks with a relatively small market value, and (ii) stocks with a relatively high book-to-market ratio. These effects are respectively referred to as the size and value anomaly.

The factor portfolios introduced by Fama & French (1993) consist of a long position in a subset of the market, combined with an equal short position in another segment, therefore resulting in a zero-investment portfolio. The size effect is captured by the Small Minus Big (SMB) factor, which contains long positions in stocks with small market values and short positions in big market value stocks. Similarly, the value effect is captured by the High Minus Low (HML) factor, which contains long positions in stocks with high book-to-value ratios and short positions in stocks for which this ratio is low.

Fama & French (1993) do not replace the market's return of the CAPM in equation (2), but augment it with the SMB and HML factors, leading to the following model:

$$E(r_i) = \beta_{M;i}E(r_M) + \beta_{SMB;i}E(r_{SMB}) + \beta_{HML;i}E(r_{HML}), \tag{3}$$

where $E(r_{SMB})$ and $E(r_{HML})$ refer to the expected returns of respectively the SMB and HML factors, with $\beta_{SMB;i}$ and $\beta_{HML;i}$ as the respective sensitivities of asset i to these factors; and the same definitions for the remaining terms as in equations (1) and (2).

Fama & French (1993) empirically compare the FF3 model to the CAPM, and conclude that the FF3 is better able to explain stock returns than the CAPM. They conclude this based on a set of timeseries regressions that is each fitted on the historical returns of a test asset, using either the CAPM or FF3. Contrary to the theoretical models (2) and (3) that do not contain intercepts, the fitted regressions allow for non-zero intercepts. These fitted intercepts then represent the pricing error of the model. The analyses by Fama & French (1993) compare factor sensitivities and pricing errors between the models, using a combination of individual and joint significance testing. The joint significance tests follow the Gibbons, Ross & Shanken (1989) test, which leads to the conclusion that both CAPM and FF3 result in significant pricing error, although the latter manages to reduce them compared to the CAPM. The factor model analyses performed in this thesis are based on those in Fama & French (1993) and are explained in more detail in chapter 4.

The FF3 factor model is followed by many other multi-factor models. As an illustration, Harvey, Liu & Zhu (2016) count 316 factors published in top journals. Many of these factors lead to reduced pricing errors in one specific empirical setting but fail to do so in other samples. The search for an asset pricing model that robustly explains returns without pricing error is therefore still an open research gap.

2.2 Sustainability Risk

Sustainability risks possibly are such a source of systematic risk that would justify a risk premium for exposed investors. That is, sustainability risk can pose both threats and opportunities to firms, that could result in both under-and overperformance compared to competitors. This in turn can help explain differences in stock returns.

Because of various societal shifts and regulatory pressure, numerous businesses are facing new and complex risks and opportunities: climate change, demographic changes, geopolitical tensions, deglobalization, and growing economic inequality (KPMG, 2019). For instance, industries heavily reliant on carbon, like oil & gas, mining, and chemicals, are experiencing disruption due to climate change regulations and the risks of stranded assets, i.e., physical risk.

Regulatory bodies in developed economies, such as Europe and the USA, increasingly require companies to proactively implement mitigating measures and comprehensively report their sustainability efforts. Sustainability reporting outlines how companies disclose their sustainability performance. It is primarily divided into three main metrics: Environmental, Social, and Governance pillars, also jointly known as ESG. The Environmental pillar focuses on compliance with environmental regulations; the Social pillar addresses fair treatment of stakeholders and the surrounding community; and the Governance pillar enhances corporate ethics, transparency, audits, internal controls, and shareholder rights (Limkriangkrai et al., 2017). This shift has led to widespread adoption of ESG disclosure by companies aiming to engage with investor demand, comply with regulations, build credibility to the public, and stay competitive (Olsen & Ensbuk, 2023).

Along with this widespread adoption, the academic community is interested in investigating whether such adoption provides a competitive advantage for companies. Regarding the effect of ESG disclosure on investment returns, a study by Lins et al. (2017) examines the relation between ESG performance and stock returns, following the 2008 financial crisis. They find that companies with strong social responsibility manage risks better and achieve higher stock returns, compared to those with poor social responsibility scores. Lioui (2018) finds that companies with good ESG performance tend to outperform during bad ESG-related conditions and underperform during good conditions. Similarly, Hong and Kacperczyk (2009) shows that companies with good ESG performance are less risky and earn lower expected returns, while companies with poor ESG performance are riskier, leading to higher capital costs for "sin" firms.

Based on these findings, firms' ESG performance is related to their (expected) profitability through various regulatory, reputational and other related risks. However, sustainability risk as an asset pricing factor is still an underexplored academic field.

2.3 Sustainability as an Asset Pricing Factor

The previous two sections provide a literature review and theoretical framework on asset pricing and sustainability risk, for which the following research gaps are respectively identified:

1. There is no academic consensus yet on a factor model that can explain stock returns without pricing error.
2. Sustainability risk affects firms' (expected) profitability, but its validity as an asset pricing factor is not yet widely examined in academic literature.

This thesis contributes to filling the intersection of these research gaps by examining whether factors models considering sustainability risk can explain stock returns better than the benchmark models CAPM and FF3.

For that purpose, I build upon the theoretical asset pricing framework as outlined in section 2.1, combined with the categorisation of sustainability risk explained in section 2.2. Using this categorisation into the Environmental, Social, and Governance risks, this thesis examines each pillar's individual relevance to asset pricing, as well as the relevance of the Total ESG Risk.

In mathematical terms, this thesis introduces the following four theoretical models as extensions of FF3:

$$E(r_i) = \beta_{M;i}E(r_M) + \beta_{SMB;i}E(r_{SMB}) + \beta_{HML;i}E(r_{HML}) + \beta_{ESG;i}E(r_{ESG}), \tag{4}$$

$$E(r_i) = \beta_{M;i}E(r_M) + \beta_{SMB;i}E(r_{SMB}) + \beta_{HML;i}E(r_{HML}) + \beta_{ENV;i}E(r_{ENV}), \tag{5}$$

$$E(r_i) = \beta_{M;i}E(r_M) + \beta_{SMB;i}E(r_{SMB}) + \beta_{HML;i}E(r_{HML}) + \beta_{SOC;i}E(r_{SOC}), \tag{6}$$

$$E(r_i) = \beta_{M;i}E(r_M) + \beta_{SMB;i}E(r_{SMB}) + \beta_{HML;i}E(r_{HML}) + \beta_{GOV;i}E(r_{GOV}), \tag{7}$$

where $E(r_{ESG})$, $E(r_{ENV})$, $E(r_{SOC})$, and $E(r_{GOV})$ correspond to expected returns of factors capturing respectively Total ESG risk, Environmental risk, Social risk, and Governance risk; $\beta_{ESG;i}$, $\beta_{ENV;i}$, β_{SOCi}, and $\beta_{GOV;i}$ indicate the respective sensitivities of asset i to these factors. These factors are constructed using a similar long-short logic as in Fama & French (1993).

The factor construction is explained in more detail in chapter 3. The remainder of this thesis uses the term 'sustainability factor models' to refer to all four theoretical models described by equations (4) – (7). The term 'sustainability factors' refers to the four factor portfolios created for Total ESG Risk and the three pillar scores.

I pose and assess three hypotheses regarding the sustainability factor models in this thesis. Considering the academic findings in section 2.2 that sustainability risk likely affects firms' performance, it is plausible that sustainability factors have explanatory power for stock returns. This leads to the first hypothesis:

H1: Some of the sustainability factors have significant explanatory power of stock returns.

Since sustainability risks are not directly considered by the benchmark factors (i.e., market, SMB, and HML), it is likely that sustainability factors add relevant information to the models that is not available through (a combination of) the benchmark factors. This leads to the second hypothesis:

H2: Some of the sustainability factor models can explain a larger proportion of stock return variation than the benchmark models CAPM and FF3.

Finally, if sustainability factors are indeed valid asset pricing factors, then their inclusion in a model would bring it closer to *the* theoretical asset pricing model that explains returns without pricing errors, i.e., intercepts equal to zero. This leads to the third and final hypothesis:

H3: Some of the sustainability factor models lead to pricing errors closer to zero than the benchmark models CAPM and FF3.

Without known academic indications in favour of any particular sustainability factor above the others, this thesis does not single out a specific sustainability factor in any of the hypotheses.

3 Data Collection and Description

To validate and compare the theoretical asset pricing models of the previous chapter, I collect and prepare a historic sample of monthly return data on the US stock market. The required data sample consists of the explanatory factor returns and the test asset returns to be explained. After preprocessing, the dataset consists of 134 months of returns from February 2013 until March 2024.

This thesis denotes the monthly return observations with a subscript t (e.g., $r_{i;t}$) to indicate that the observed timeseries is meant. The notation without this subscript (e.g., r_i) refers to the random variable from which the observations are drawn.

3.1 Factor Returns

3.1.1 Benchmark Factors

The monthly return data of the benchmark factors Market ($r_{M;t}$), SMB ($r_{SMB;t}$), and HML ($r_{HML;t}$) are obtained from the publicly available Fama-French Data Library[2]. This dataset is managed by Kenneth R. French and regularly updated with the latest return data. The factor construction method is explained in detail in Fama & French (2023), which is mostly unchanged compared to the original method of Fama & French (1993).

This dataset also provides the monthly risk-free rates of return, which are used to calculate excess returns from total returns in this thesis. The risk-free rates are based on the one-month US Treasury bill rates.

3.1.2 Sustainability Factors

I construct the four sustainability factor portfolios based on the S&P 500 index and data from Refinitiv Eikon. For every month in the period from January 2013 – March 2024, I retrieve the variables in Table 1 using the Refinitiv Data Library API for Python.

[2] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Table 1
Variables from Refinitiv Eikon

Variable	Description
$PriceClose_{i;t}$	Last closing price of stock i in month t
$IssueMarketCap_{i;t}$	Last total market value of stock i in month t
$TRESGCScore_{i;t}$	Latest total ESG score of stock i published by month t
$EnvironmentPillarScore_{i;t}$	Latest Environment score of stock i published by month t
$SocialPillarScore_{i;t}$	Latest Social score of stock i published by month t
$GovernancePillarScore_{i;t}$	Latest Governance score of stock i published by month t

The variable names and descriptions of all variables extracted from Refinitiv Eikon. The variable names are equal to Refinitiv's data item codes without the "TR." in front. The data item codes uniquely identify variables in Refinitiv.

The total return of each stock i in month t is calculated as the percentage change in closing price compared to month $t - 1$. The excess return $r_{i;t}$ is then calculated as the total return minus the risk-free rate during month t.

The total ESG score and the three pillars' scores range from 0 to 100, with higher scores indicating better performance on the pillars. The scores are updated weekly by LSEG – the parent company of Refinitiv. LSEG's scoring methodology[3] uses 186 comparable firm metrics from publicly available data to compare each firm's ESG-performance to similar firms in its peer group. The three pillar scores are a weighted sum of the underlying metrics. Industry-specific weights are applied to the Environment and Social pillars; Governance weights are equal across industries. The total ESG score consists of a weighted sum of the three pillar scores, adjusted for any recent ESG controversies, i.e., recent negative ESG-related media coverage on the firm.

The factor returns $r_{ESG;t}$, $r_{ENV;t}$, $r_{SOC;t}$, and $r_{GOV;t}$ are constructed based on the corresponding four sustainability variables. The portfolios are constructed as value-weighted long-short portfolios. At the start of each month t, all stocks are ranked on the four sustainability scores as known by the end of month $t - 1$. The long leg of the portfolio consists of one third of the stocks with the best (i.e., highest) scores in the previous month;

[3] For a more detailed description of the methodology, see https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf.

the short leg consists of one third of the worst (i.e., lowest) scores in the previous month. The stocks within each of the two legs are weighted based on the stocks' market value in month $t - 1$. The total of the long positions is equal to the total of the short positions, resulting in a zero-investment portfolio. The factor's return is then equal to the return of the long-short portfolio during month t.

If the observation of any of the variables in Table 1 is missing for a stock i in month t and/or $t - 1$, then this stock is omitted from all factors' construction in month t. This implies that the factor timeseries start in the second month of collected data, i.e., February 2013.

3.1.3 Descriptive Statistics Factors

Table 2 provides some descriptive statistics on the excess returns of the seven factor portfolios. One of the most apparent observations is that only the Market factor has a positive average return during the sample period (February 2013 – March 2024). Therefore, if the factor portfolios are purely regarded as an investment, only the Market would have led to a favourable outcome during this period. However, average factor returns do not directly indicate whether the portfolios provide explanatory power in an asset pricing setting. Furthermore, the Market outperforming other factor is not unusual; the average returns of SMB and HML in the sample of Fama & French (1993) are also lower than the Market.

Table 2
Factor Return Descriptive Statistics

Factor	Mean	Std.	Min.	Max.
Market	1.08	4.39	-13.39	13.65
SMB	-0.12	2.68	-5.94	7.36
HML	-0.13	3.58	-13.87	12.75
Total ESG Risk	-0.42	2.50	-7.41	5.20
Environmental Risk	-0.25	1.75	-4.74	6.28
Social Risk	-0.14	1.70	-4.66	5.23
Governance Risk	-0.03	1.71	-4.71	4.88

Descriptive statistics for the factor portfolios Market ($r_{M;t}$), SMB ($r_{SMB;t}$), HML ($r_{HML;t}$), Total ESG Risk ($r_{ESG;t}$), Environmental Risk ($r_{ENV;t}$), Social Risk ($r_{SOC;t}$), and Governance Risk ($r_{GOV;t}$). All numbers are in percentages (%). The sample consists of 134 monthly returns in the period February 2013 – March 2024.

Correlations between the seven factor return timeseries are presented in Table 3. The benchmark factors are weakly correlated with the sustainability factors. This can be seen from the correlation coefficients between any of the benchmark factors with any of the sustainability factors, which are all below 0.50 in absolute value. This can be interpreted as the sustainability factors containing different return information than the benchmark factors. However, whether this information makes them suitable explanatory variables (i.e., asset pricing factors) cannot be concluded based on this. Nevertheless, it shows that adding any of the sustainability factors to the benchmark models would not cause multicollinearity. Hence, it is worthwhile to investigate the sustainability factor models in equations (4) – (7).

Table 3
Correlations between factor returns

	Market	SMB	HML	ESG	Env.	Soc.	Gov.
Market	1.00						
SMB	0.30	1.00					
HML	0.04	0.06	1.00				
ESG	-0.25	-0.01	0.46	1.00			
Env.	-0.12	-0.29	0.04	0.12	1.00		
Soc.	-0.15	-0.21	-0.08	0.08	0.78	1.00	
Gov.	0.12	-0.05	0.02	-0.01	0.53	0.60	1.00

The Pearson correlation coefficients between the factor portfolios Market ($r_{M;t}$), SMB ($r_{SMB;t}$), HML ($r_{HML;t}$), Total ESG Risk ($r_{ESG;t}$), Environmental Risk ($r_{ENV;t}$), Social Risk ($r_{SOC;t}$), and Governance Risk ($r_{GOV;t}$). The sample consists of 134 monthly returns in the period February 2013 – March 2024.

The correlations between the three sustainability factors based on a single pillar – i.e., Environmental Risk ($r_{ENV;t}$), Social Risk ($r_{SOC;t}$), and Governance Risk ($r_{GOV;t}$) – are relatively high. However, the correlations between these three factors and the Total ESG Risk factor are considerably lower. This indicates that the adjustment for ESG controversies to the total ESG score tends to have a large effect on the final score compared to the underlying pillar scores, which results in considerably different portfolio compositions and returns.

3.2 Test Asset Returns

The factor portfolios constructed in section 3.1 are to be assessed for their ability to explain the returns of a set of test assets. These test assets are 25 double-sorted portfolios on size (i.e.

market value) and book-to-market ratio. The monthly returns are retrieved from the Fama-French Data Library. Although there are several caveats that I explain in this section, I use these test assets because they are (i) suitable for my testing methodology, (ii) widely applied, and (iii) less prone to some biases compared to a number of alternatives.

In theory, an asset pricing model must be able to explain the excess return of any individual stock. However, individual stocks' return data is often regarded as too noisy to serve as test assets (Jensen et al., 1972; Fama & MacBeth, 1973). Furthermore, using individual stocks leads to a large cross-sectional dimension compared to the timeseries dimension, i.e., the number of test assets would be large compared to the number of observations over time. This causes statistical issues that some of the most widely used asset pricing tests are not able to handle, such as the test by Gibbons, Ross & Shanken (1989). Therefore, most asset pricing studies use characteristic-sorted portfolios as test assets.

The use of characteristic-sorted portfolios as test assets is also not free from critique itself. Mainly, these portfolios are likely to cause a bias in favour of factor portfolios based on the same characteristics as the test assets (Ferson & Harvey, 1999; Berk, 2000; Pukthuanthong, Roll & Subrahmanyam, 2019). That is, factor portfolios tend to be better at explaining the returns of test assets based on the same characteristics.

Although any bias is preferably avoided, I argue that it is more reasonable to have a potential bias working in favour of the benchmark factors being challenged, rather than in favour of the "challenger" sustainability factors. I therefore use test assets sorted by size and value – the same characteristics as in SMB and HML – instead of total ESG score or its pillar scores.

4 Analysis and Findings

I use timeseries regressions to empirically compare the four sustainability factor models – i.e., equations (4) – (7) – with the benchmark models CAPM and FF3. All six factor models are fitted to explain each of the 25 timeseries of test asset returns, resulting in 150 regression models to fit. The results are then compared to check each of the three hypotheses H1 – H3. The analyses are based on those of Fama & French (1993).

None of the theoretical asset pricing models in chapter 2 contain intercepts. In the empirical tests, however, intercepts are included. The estimated intercepts in the regressions show how much of the observed test asset's return is left unexplained by the model on average. These estimated intercepts are therefore referred to as pricing errors.

For example, to empirically assess the sustainability factor model based on total ESG risk (i.e., FF3 + r_{ESG}), equation (4) from chapter 2 needs to be rewritten into a testable timeseries model for each of the 25 test assets. Specifically, I rewrite the theoretical model

$$E(r_i) = \beta_{M;i}E(r_M) + \beta_{SMB;i}E(r_{SMB}) + \beta_{HML;i}E(r_{HML}) + \beta_{ESG;i}E(r_{ESG}), \tag{4}$$

such that, for test asset $i = 1, \dots, 25$, it becomes:

$$r_{i;t} = \alpha_i + \beta_{M;i}r_{M;t} + \beta_{SMB;i}r_{SMB;t} + \beta_{HML;i}r_{HML;t} + \beta_{ESG;i}r_{ESG;t} + e_{i;t}, \tag{8}$$

where $r_{i;t}$ is the observed return of test asset i in month t; α_i is the average pricing error for test asset i; and $e_{i;t}$ is the regression residual of test asset i in month t. The other five theoretical factor models are rewritten into regression models following the same steps.

Each of the 125 timeseries regressions is fitted using Ordinary Least Squares (OLS). The following three subsections sequentially assess hypotheses H1 – H3 using the regression results. The main text of this thesis only provides the aggregated results that are necessary for this purpose. The full underlying results per fitted timeseries regression are available in tables Table A-1 – Table A-6 in Appendix A.

4.1 Explanatory Power of Sustainability Factors (Hypothesis H1)

To assess the first hypothesis, I check whether any of the four sustainability factors provide significant explanatory power. That is, I check the size and significance of the estimates of $\beta_{ESG;i}$, $\beta_{ENV;i}$, $\beta_{SOC;i}$, and $\beta_{GOV;i}$ if they each get individually added to the FF3 model. The significance is determined using the p-values at various significance levels.

Table 4 presents a summary of the estimated factor sensitivities. The factor returns based on Environmental Risk ($r_{ENV;t}$) and Social Risk ($r_{SOC;t}$) have a significant explanatory power of more than half of the 25 size-value test portfolios, at the common significance level of 5%. For the other two sustainability factors, the number of significant sensitivities is substantially lower. In comparison, almost all estimated factor sensitives for the Market, SMB, and HML are significant at the 1% level in all model specifications (see Appendix A).

Table 4
Summary of estimated sensitivities to sustainability factors added to FF3

	β-estimates			#Significant at level (out of 25)		
	Mean	Min.	Max.	1%	5%	10%
$\beta_{ESG;i}$	0.033	-0.210	0.260	3	7	10
$\beta_{ENV;i}$	-0.163	-0.474	0.623	13	17	17
$\beta_{SOC;i}$	-0.100	-0.326	0.535	13	14	14
$\beta_{GOV;i}$	-0.062	-0.223	0.169	4	7	10

Summary statistics and number of significant estimates of sensitivity to the four sustainability factors. Each row represents a different model, in which the respective factor is added as fourth factor to the FF3 benchmark model. The full results are provided in Appendix A.

The β-estimates of all four sustainability factors range from negative to positive numbers, but only the estimate of $\beta_{ESG;i}$ is positive on average. This means that the average test asset tends to move in the same direction as the Total ESG Risk factor, while it tends to move in the opposite direction for the other factors.

In conclusion, there are mixed indications regarding hypothesis H1. Although all four sustainability factors have significant explanatory power for some of the test assets, the share of significant betas is substantially lower than for benchmark factors.

4.2 Percentage of Explained Return Variation (Hypothesis H2)

The percentage of variation in the returns of the test assets that a model can explain is commonly measured by the model's R^2. If this metric is 1, then the explanatory variables – i.e., the factors – can perfectly explain the variation in the variable to be explained – i.e., the test asset returns. A major caveat of the R^2-metric is that simply adding another factor to a model would always results in an R^2 that is at least as high as that of the smaller model. This means that the R^2 of all sustainability factor models cannot be worse than those of the benchmark models CAPM and FF3, only because they contain more factors.

Therefore, the normal R^2 is unsuitable to assess hypothesis H2 fairly. As an alternative, I use the adjusted R^2 metric, which accounts for the aforementioned limitation of the normal R^2. The adjusted R^2 is calculated as follows:

$$R^2_{adj} = 1 - (1 - R^2)\frac{T - 1}{T - K - 1}, \tag{9}$$

where T is the sample size and K is the number of factors. Hence, in this case $T = 134$, i.e., the number of months in the timeseries data.

Table 5 presents the average adjusted R^2 over the 25 size-value test assets per sustainability factor model. On average, all four factor models explain a larger proportion of the test assets' return variation than both the CAPM and FF3. The increase in average R^2 from CAPM to FF3 is substantially larger than the increase from FF3 to any of the sustainability factor models. Considering that adding the SMB and HML factor to the CAPM (i.e., FF3) increased the average explained variation from 73.1% to 91.4%, it would be mathematically impossible for any additional factor to result in a similar increase.

Table 5
Adjusted R^2 comparison

Factor model	Mean R^2_{adj}	#Higher than benchmark	
		CAPM (avg.=0.731)	FF3 (avg.=0.914)
FF3 + Total ESG Risk	0.923	25	23
FF3 + Environmental Risk	0.927	25	23
FF3 + Social Risk	0.925	25	23
FF3 + Governance Risk	0.922	25	23

The average adjusted R^2 is calculated over the 25 individual adjusted R^2 metrics that are calculated over the timeseries regressions of the size-value portfolios on the factor models. The comparison with the benchmark models indicates how many of the 25 adjusted R^2 metrics are higher for the sustainability factor model (row) than for the benchmark (column).

Not only on average, also individually all of the 25 adjusted R^2 metrics are higher for the sustainability factor models than for the CAPM. Only two size-value test assets have a higher adjusted R^2 in the FF3 model than in the four sustainability factor models.

In conclusion on hypothesis H2, there is a strong indication that all four sustainability factor models explain a larger proportion of return variation than the CAPM. There is also indication that this proportion is also larger compared to the FF3, but the difference is less substantial.

4.3 Pricing Errors (Hypothesis H3)

The assessment of hypothesis H1 shows that the sustainability factors have a significant explanatory power of the test assets' excess returns. Furthermore, the analysis of hypothesis H2 shows that adding either of the four sustainability factors to FF3 increases the ability of the overall model to explain the return variation in the test assets. What remains to be assessed is whether any of the four sustainability factor models are sufficient to describe the cross-section of stock returns. That is, I assess whether any of these models succeed at explaining the 25 test asset's returns – the proxy for the full cross-section – with a pricing error indistinguishable from zero.

If an asset pricing factor model explains any single asset's return without significant pricing error, then this is no sufficient evidence to claim it can explain the full cross-section of assets. Vice versa, if the model explains a single asset with significant pricing errors, then this does not directly invalidate the model for the overall cross-section of assets. Hence, the pricing errors for the set of test assets need to be *jointly* insignificant.

To test the joint significance of the estimated intercepts, I use the statistical test of Gibbons, Ross & Shanken (1989), commonly referred to as the GRS test. This is the same test as used by Fama & French (1993) to assess the FF3 model. The test is a form of an F-test and assesses the following null hypothesis against its alternative:

H_0: $[\alpha_1 \quad \cdots \quad \alpha_{25}] = \vec{0}$,

H_1: *H_0 is not true,*

where $[\alpha_1 \quad \cdots \quad \alpha_{25}]$ is the vector of estimated intercepts and $\vec{0}$ is a vector of zeros. The GRS test statistic is calculated over the estimated intercepts and covariance matrix of residuals. The test statistic follows an F-distribution, under the null hypothesis and the assumption that all returns follow a normal distribution Gibbons, Ross & Shanken (1989). The degrees of freedom of the F-distribution depend on the number of observations and factors in the estimated models. Using this distribution, p-values are calculated.

Table 6 presents the results of the GRS test for the benchmark and sustainability factor models. Besides the model incorporating Environmental Risk, all factor models have p-values above the 5% significance level. These non-significant GRS statistics indicate that the null hypothesis cannot be rejected for the respective models and that, hence, the models' pricing errors are jointly indistinguishable from zero.

Table 6
GRS test results

Factor model	GRS stat.	p-value
CAPM	1.296	0.181
FF3	1.276	0.196
Total ESG Risk (FF3 + r_{ESG})	1.252	0.214
Environmental Risk (FF3 + r_{ENV})	2.030	0.007
Social Risk (FF3 + r_{SOC})	1.606	0.051
Governance Risk (FF3 + r_{GOV})	1.394	0.125

GRS test statistics and p-values for the null hypothesis that the 25 estimated intercepts (i.e., pricing errors) are jointly equal to zero.

Although these results seem to confirm the validity of five out of six factor models, a cautionary note is necessary. Note that even the pricing errors of the CAPM are not significantly different from zero, while numerous other studies note significant pricing errors for this model. For example, Fama & French (1993) conclude that both the CAPM and FF3 result in significant pricing errors. Their research and this thesis both use the same 25 test assets; the same Market, SMB, and HML factor specifications; and identical GRS test calculations. Therefore, the most likely explanations for the difference in results are the differences in sample periods and sizes.

The sample period of this thesis (i.e., February 2013 – March 2024) has a large time gap with the period in Fama & French (1993), which spanned July 1963 – December 1991. The specific (financial) economic circumstances during this thesis' sample period might result in stock returns that are more easily described by the Market, SMB, and HML factors.

Besides the different sampled periods in time, a more likely explanation is the difference in sample sizes. While this thesis uses timeseries regressions on 134 months of data, Fama & French (1993) use 342 months. The smaller timeseries dimension in this thesis can negatively affect the power of the GRS test – i.e., the ability of the test to reject a false null hypothesis (Gibbons, Ross & Shanken, 1989).

Regardless of the statistical significance of the GRS test statistics, the value of statistic is often used as a heuristic to compare the performance of asset pricing models (e.g., Fama & French 1993; Fama & French, 2015a). That is, a smaller value of the GRS statistic is interpreted as the model being better at reducing the price errors towards zero, even if those pricing errors are significantly different from zero.

Following this heuristic interpretation, Table 6 shows that the model incorporating Total ESG Risk (FF3 + r_{ESG}) results in the smallest GRS statistic. This implies that this model performs better than both the CAPM and FF3 benchmark models, in terms of reducing pricing errors. However, the three sustainability factor models using pillar scores result in higher GRS statistics than the benchmarks.

In conclusion on hypothesis H3, one of the sustainability factor models indeed manages to explain the cross-section of asset returns with smaller pricing errors than the benchmark models. The Total ESG Risk factor can therefore be considered a valuable asset pricing factor, at least for the examined sample. The other three sustainability factors do not seem to improve the benchmark models CAPM and FF3.

5 Discussion and Conclusion

5.1 Academic Implications

The three assessed hypotheses each provide relevant insights on the central question whether sustainability factors can augment the FF3 model such that it can more accurately explain stock returns. Although all three are relevant, the last one is the most essential to address the research question. That is, significant sensitivities to the sustainability factors (H1) can indicate that the factors are useful to include in the model, but this does not necessarily mean that the overall model becomes better at explaining returns. The comparison of the adjusted R^2-metrics (H2) provides better insights in the overall model performance, but it does not statistically test the models' validity. Such a formal test is provided by the GRS test (H3), which is therefore the most important analysis to answer the research question.

The lowest GRS test statistic is achieved by the asset pricing model that adds Total ESG Risk to the Market, SMB, and HML. This is also the only model that manages to reduce pricing errors more than the benchmarks CAPM and FF3. Although the results on the Total ESG Risk seem affirmative for the research question, they also raise follow-up questions.

A question to be answered in future research is why the Total ESG Risk factor reduces pricing errors, while the factors based on its three underlying pillars increase the pricing errors. The Total ESG Risk score is of course based on the three pillar scores, but also takes into account ESG controversies. A potential explanation could therefore be that these ESG controversies contain the information that is most explanatory for returns. An alternative explanation could be that the success of Total ESG Risk is mostly due to its underlying pillar scores – rather than the ESG controversies – but that a pillar's score is only informative in combination with the context of the other two, therefore only making the Total ESG Risk a valid factor. These two explanations could be assessed in future research by constructing another Total ESG Risk factor that is based on the three pillar scores but excludes ESG controversies. The performance of that factor could then be compared to the one in this thesis.

5.2 Implications for Finance Practitioners

The results imply an important implication for finance practitioners: asset pricing models that do not consider Total ESG Risk seem to omit a factor. This can affect, e.g., the performance assessment of investment managers, because it is often measured as their achieved return on top what is expected according to the used asset pricing model. An omitted factor might assign overperformance to the investment manager's skills in picking the right stocks, while it is in fact due to exposure to the omitted factor. Another example of the practical implications of the factor's omission is the determination of the WACC within corporate finance. That is, the omission could lead to over- or underestimation of the WACC, which could result in over- or undervaluation of a firm. To avoid these consequences, finance practitioners should consider measuring their exposure to the Total ESG Risk factor.

5.3 Limitations

This thesis comes with several limitations, related to the available data, the scope, and testing methodology. Regarding the data, this thesis uses a sample size that is relatively limited (i.e., 134 months), compared to e.g., Fama & French (1993) who use 342 months of data. The limited sample period is mostly caused by a practical limitation: there is not as much high-quality historic data on sustainability metrics available as there is on stock prices. The relatively short timeseries might have led to the inability to reject the pricing errors in the GRS tests.

The scope of the analyses in this thesis is limited to the S&P 500 index. Ideally, a factor is able to explain returns in various markets and times. Future research is needed to verify whether the Total ESG Risk factor can also explain returns in different geographies and time periods.

Regarding the testing methodology, the use of the GRS test inherently implies a few limitations. First, the test statistic cannot be calculated if the cross-sectional dimension is larger than the timeseries dimension, therefore requiring the use of portfolios as test assets

(Gibbons, Ross & Shanken, 1989). Second, the GRS test assumes the returns to be normally distributed. An alternative testing framework that could remove these two limitations on the GRS test is proposed by Harvey & Liu (2021). I therefore suggest future research to apply this alternative testing approach to the sustainability factor models assessed in this thesis.

5.4 Conclusion

This thesis provides empirical indications that an asset pricing factor based on Total ESG Risk – when added to the FF3 model – can explain stock returns with lower pricing errors than the CAPM and FF3. Although the results provide an affirmative answer to the research question, the GRS test results indicate that the power of the test might not be high enough. Both academic researchers and finance practitioners are therefore encouraged to further examine the validity of this factor, including in different geographical and temporal contexts.

References

Berk, J. B. (2000). Sorting out sorts. *The Journal of Finance, 55*(1), 407-427.

Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance, 56*(3), 921-965.

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

Fama, E. F., & French, K. R. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics,* 166(1), 1-22.

Fama, E.F., & French, K.R. (2023). Production of U.S. Rm-Rf, SMB, HML in the Fama-French Data Library. *Chicago Both Paper,* No 22-23.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*(3), 607-636.

Ferson, W. E., & Harvey, C. R. (1999). Conditioning variables and the cross section of stock returns. *The Journal of Finance, 54*(4), 1325-1360.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society,* , 1121-1152.

Harvey, C. R., & Liu, Y. (2021). Lucky factors. *Journal of Financial Economics, 141*(2), 413-435.

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

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

Jensen, M. C., Black, F., & Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests.

KPMG Internationaal. (2019). Winning Strategies for Long-Term: How to Create Value and Enhance Competitiveness in the Age of Distruption and Short-termism. *Amstelveen, The Netherlands.*

Limkriangkrai, M., Koh, S., & Durand, R. B. (2017). Environmental, social, and governance (ESG) profiles, stock returns, and financial policy: Australian evidence. *International Review of Finance, 17*(3), 461-471.

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

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance, 20*(4), 587-615.

Lioui, A. (2018). Is ESG risk priced? *Available at SSRN 3285091,*

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance, 7*(1), 77-91. 10.2307/2975974

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society,* , 768-783.

Olsen, C., & Enbusk, H. (2023). No title. *Can Private Equity Enhance Value Creation by Integrating ESG Considerations Throughout the Investment Life Cycle?,*

Pukthuanthong, K., Roll, R., & Subrahmanyam, A. (2019). A protocol for factor identification. *The Review of Financial Studies, 32*(4), 1573-1607.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442.

Treynor, J. L. (1961). Market value, time, and risk. *Time, and Risk (August 8, 1961),*

Appendix A Regression Results

Table A-1
CAPM regression results

$$r_{i;t} = \alpha_i + \beta_{M;i} r_{M;t} + e_{i;t}$$

	Low	2	3	4	High	Low	2	3	4	High
			α_i (in %)					p-value α_i		
Small	-0.847	-0.457	-0.392	-0.327	-0.004	0.070	0.222	0.208	0.311	0.994
2	-0.449	-0.176	-0.334	-0.287	-0.448	0.199	0.492	0.182	0.340	0.227
3	-0.405	-0.138	-0.302	-0.312	-0.502	0.120	0.449	0.138	0.239	0.185
4	-0.077	-0.083	-0.111	-0.378	-0.347	0.662	0.533	0.545	0.136	0.326
Big	0.297	0.058	-0.012	-0.199	-0.181	0.037	0.614	0.941	0.389	0.593
			$\beta_{M;i}$					p-value $\beta_{M;i}$		
Small	1.320	1.197	1.133	1.087	1.085	0.000	0.000	0.000	0.000	0.000
2	1.304	1.177	1.183	1.111	1.279	0.000	0.000	0.000	0.000	0.000
3	1.192	1.126	1.089	1.193	1.300	0.000	0.000	0.000	0.000	0.000
4	1.092	1.095	1.082	1.170	1.195	0.000	0.000	0.000	0.000	0.000
Big	1.001	0.855	0.882	0.935	1.220	0.000	0.000	0.000	0.000	0.000
			Adj. R^2							
Small	0.551	0.610	0.669	0.635	0.438					
2	0.680	0.762	0.775	0.677	0.646					
3	0.763	0.853	0.815	0.756	0.643					
4	0.855	0.913	0.843	0.766	0.636					
Big	0.885	0.895	0.824	0.715	0.663					

Table A-2
FF3 regression results

$$r_{i;t} = \alpha_i + \beta_{M;i} r_{M;t} + \beta_{SMB;i} r_{SMB;t} + \beta_{HML;i} r_{HML;t} + e_{i;t}$$

	Low	2	3	4	High	Low	2	3	4	High
			α_i (in %)					p-value α_i		
Small	-0.598	-0.256	-0.216	-0.124	0.301	0.032	0.239	0.146	0.286	0.334
2	-0.342	-0.097	-0.278	-0.163	-0.287	0.081	0.445	0.013	0.161	0.055
3	-0.406	-0.178	-0.330	-0.289	-0.389	0.022	0.145	0.007	0.018	0.011
4	-0.227	-0.236	-0.230	-0.449	-0.332	0.127	0.057	0.091	0.012	0.070
Big	-0.056	-0.206	-0.217	-0.371	-0.253	0.465	0.074	0.107	0.002	0.231
			$\beta_{M;i}$					p-value $\beta_{M;i}$		
Small	1.025	0.960	0.928	0.891	0.846	0.000	0.000	0.000	0.000	0.000
2	1.079	1.009	1.033	0.952	1.106	0.000	0.000	0.000	0.000	0.000
3	1.043	1.024	1.000	1.075	1.142	0.000	0.000	0.000	0.000	0.000
4	1.024	1.056	1.041	1.094	1.089	0.000	0.000	0.000	0.000	0.000
Big	1.065	0.895	0.903	0.948	1.162	0.000	0.000	0.000	0.000	0.000
			$\beta_{SMB;i}$					p-value $\beta_{SMB;i}$		
Small	1.851	1.534	1.316	1.203	1.405	0.000	0.000	0.000	0.000	0.000
2	1.469	1.165	1.046	1.031	1.080	0.000	0.000	0.000	0.000	0.000
3	1.040	0.778	0.680	0.790	0.968	0.000	0.000	0.000	0.000	0.000
4	0.586	0.418	0.397	0.550	0.595	0.000	0.000	0.000	0.000	0.000
Big	-0.002	-0.021	0.045	0.031	0.377	0.932	0.612	0.355	0.481	0.000
			$\beta_{HML;i}$					p-value $\beta_{HML;i}$		
Small	-0.247	-0.069	0.226	0.547	0.709	0.001	0.234	0.000	0.000	0.000
2	-0.285	0.039	0.283	0.574	0.844	0.000	0.255	0.000	0.000	0.000
3	-0.269	0.116	0.356	0.599	0.931	0.000	0.000	0.000	0.000	0.000
4	-0.256	0.145	0.385	0.533	0.933	0.000	0.000	0.000	0.000	0.000
Big	-0.336	0.039	0.302	0.636	0.878	0.000	0.209	0.000	0.000	0.000
			Adj. R^2							
Small	0.850	0.875	0.929	0.955	0.775					
2	0.905	0.945	0.958	0.954	0.946					
3	0.898	0.938	0.938	0.951	0.945					
4	0.902	0.928	0.918	0.891	0.907					
Big	0.968	0.899	0.881	0.927	0.877					

Table A-3
Total ESG Risk factor model regression results

$$r_{i;t} = \alpha_i + \beta_{M;i} r_{M;t} + \beta_{SMB;i} r_{SMB;t} + \beta_{HML;i} r_{HML;t} + \beta_{ESG;i} r_{ESG;t} + e_{i;t}$$

	Low	2	3	4	High	Low	2	3	4	High
			α_i (in %)					p-value α_i		
Small	-0.361	-0.001	0.019	0.085	0.510	0.175	0.998	0.895	0.444	0.098
2	-0.073	0.128	-0.022	0.079	-0.047	0.680	0.289	0.849	0.509	0.732
3	-0.188	0.074	-0.064	-0.024	-0.137	0.226	0.519	0.586	0.832	0.339
4	0.020	0.036	0.036	-0.207	-0.109	0.877	0.745	0.778	0.220	0.503
Big	0.167	0.029	0.050	-0.109	-0.036	0.020	0.788	0.676	0.314	0.852
			$\beta_{M;i}$					p-value $\beta_{M;i}$		
Small	1.031	0.981	0.929	0.880	0.844	0.000	0.000	0.000	0.000	0.000
2	1.103	0.987	1.017	0.941	1.082	0.000	0.000	0.000	0.000	0.000
3	1.038	1.020	1.004	1.072	1.118	0.000	0.000	0.000	0.000	0.000
4	1.039	1.066	1.046	1.075	1.074	0.000	0.000	0.000	0.000	0.000
Big	1.021	0.902	0.932	0.967	1.128	0.000	0.000	0.000	0.000	0.000
			$\beta_{SMB;i}$					p-value $\beta_{SMB;i}$		
Small	1.674	1.365	1.159	1.055	1.262	0.000	0.000	0.000	0.000	0.000
2	1.277	0.999	0.877	0.874	0.898	0.000	0.000	0.000	0.000	0.000
3	0.861	0.605	0.510	0.608	0.780	0.000	0.000	0.000	0.000	0.000
4	0.406	0.236	0.220	0.369	0.411	0.000	0.000	0.000	0.000	0.000
Big	-0.171	-0.174	-0.114	-0.134	0.186	0.000	0.000	0.013	0.001	0.011
			$\beta_{HML;i}$					p-value $\beta_{HML;i}$		
Small	-0.266	-0.131	0.196	0.544	0.688	0.001	0.041	0.000	0.000	0.000
2	-0.341	0.054	0.270	0.554	0.867	0.000	0.147	0.000	0.000	0.000
3	-0.257	0.094	0.310	0.575	0.950	0.000	0.009	0.000	0.000	0.000
4	-0.292	0.098	0.346	0.547	0.958	0.000	0.005	0.000	0.000	0.000
Big	-0.287	-0.004	0.212	0.569	0.940	0.000	0.898	0.000	0.000	0.000
			$\beta_{ESG;i}$					p-value $\beta_{ESG;i}$		
Small	0.039	0.172	0.075	-0.005	0.049	0.749	0.069	0.251	0.925	0.725
2	0.150	-0.064	0.020	0.044	-0.087	0.064	0.246	0.694	0.423	0.167
3	-0.056	0.048	0.125	0.055	-0.079	0.430	0.362	0.021	0.288	0.228
4	0.092	0.124	0.101	-0.059	-0.096	0.122	0.016	0.082	0.441	0.199
Big	-0.165	0.115	0.260	0.190	-0.210	0.000	0.021	0.000	0.000	0.017
			Adj. R^2							
Small	0.858	0.884	0.932	0.958	0.776					
2	0.921	0.949	0.954	0.950	0.952					
3	0.918	0.944	0.940	0.956	0.950					
4	0.924	0.940	0.926	0.899	0.924					
Big	0.972	0.908	0.902	0.939	0.897					

Table A-4:
Environmental Risk factor model regression results

$$r_{i;t} = \alpha_i + \beta_{M;i} r_{M;t} + \beta_{SMB;i} r_{SMB;t} + \beta_{HML;i} r_{HML;t} + \beta_{ENV;i} r_{ENV;t} + e_{i;t}$$

	Low	2	3	4	High	Low	2	3	4	High
			α_i (in %)					p-value α_i		
Small	-0.447	-0.048	-0.019	0.091	0.656	0.090	0.819	0.896	0.419	0.027
2	-0.190	0.114	-0.002	0.072	-0.046	0.269	0.344	0.988	0.550	0.741
3	-0.276	-0.002	-0.130	-0.084	-0.226	0.057	0.984	0.273	0.446	0.083
4	-0.117	-0.055	-0.071	-0.279	-0.183	0.284	0.605	0.545	0.086	0.241
Big	0.237	-0.006	-0.045	-0.189	-0.083	0.002	0.960	0.724	0.093	0.659
			$\beta_{M;i}$					p-value $\beta_{M;i}$		
Small	1.019	0.953	0.915	0.881	0.847	0.000	0.000	0.000	0.000	0.000
2	1.073	0.995	1.015	0.934	1.095	0.000	0.000	0.000	0.000	0.000
3	1.040	1.007	0.982	1.060	1.123	0.000	0.000	0.000	0.000	0.000
4	1.016	1.041	1.023	1.078	1.082	0.000	0.000	0.000	0.000	0.000
Big	1.050	0.883	0.888	0.934	1.155	0.000	0.000	0.000	0.000	0.000
			$\beta_{SMB;i}$					p-value $\beta_{SMB;i}$		
Small	1.619	1.362	1.146	1.058	1.375	0.000	0.000	0.000	0.000	0.000
2	1.220	0.978	0.895	0.877	0.883	0.000	0.000	0.000	0.000	0.000
3	0.787	0.559	0.486	0.575	0.701	0.000	0.000	0.000	0.000	0.000
4	0.324	0.193	0.161	0.306	0.341	0.000	0.000	0.001	0.000	0.000
Big	-0.151	-0.178	-0.136	-0.157	0.113	0.000	0.000	0.007	0.000	0.122
			$\beta_{HML;i}$					p-value $\beta_{HML;i}$		
Small	-0.245	-0.073	0.223	0.542	0.687	0.001	0.197	0.000	0.000	0.000
2	-0.282	0.036	0.274	0.568	0.840	0.000	0.269	0.000	0.000	0.000
3	-0.265	0.117	0.355	0.598	0.936	0.000	0.000	0.000	0.000	0.000
4	-0.249	0.147	0.389	0.536	0.937	0.000	0.000	0.000	0.000	0.000
Big	-0.346	0.035	0.302	0.636	0.881	0.000	0.244	0.000	0.000	0.000
			$\beta_{ENV;i}$					p-value $\beta_{ENV;i}$		
Small	-0.312	-0.050	-0.090	0.018	0.623	0.040	0.676	0.277	0.782	0.000
2	-0.348	-0.106	0.097	0.008	-0.067	0.001	0.126	0.139	0.904	0.405
3	-0.397	-0.266	-0.161	-0.196	-0.420	0.000	0.000	0.018	0.002	0.000
4	-0.474	-0.267	-0.346	-0.336	-0.371	0.000	0.000	0.000	0.000	0.000
Big	0.145	-0.046	-0.173	-0.166	-0.360	0.001	0.470	0.020	0.011	0.001
			Adj. R^2							
Small	0.862	0.882	0.932	0.958	0.797					
2	0.926	0.950	0.955	0.950	0.952					
3	0.930	0.950	0.940	0.959	0.960					
4	0.946	0.945	0.937	0.908	0.932					
Big	0.969	0.905	0.889	0.935	0.901					

Table A-5
Social Risk factor model regression results

$$r_{i;t} = \alpha_i + \beta_{M;i} r_{M;t} + \beta_{SMB;i} r_{SMB;t} + \beta_{HML;i} r_{HML;t} + \beta_{SOC;i} r_{SOC;t} + e_{i;t}$$

	Low	2	3	4	High	Low	2	3	4	High
			α_i (in %)					p-value α_i		
Small	-0.378	-0.018	0.004	0.093	0.560	0.155	0.931	0.980	0.404	0.060
2	-0.122	0.137	-0.017	0.072	-0.038	0.491	0.257	0.882	0.544	0.785
3	-0.208	0.035	-0.108	-0.057	-0.154	0.167	0.747	0.359	0.604	0.258
4	-0.035	-0.021	-0.021	-0.224	-0.120	0.773	0.841	0.857	0.170	0.452
Big	0.219	0.003	-0.024	-0.169	0.000	0.004	0.980	0.847	0.125	0.998
			$\beta_{M;i}$					p-value $\beta_{M;i}$		
Small	1.021	0.960	0.917	0.883	0.857	0.000	0.000	0.000	0.000	0.000
2	1.073	0.996	1.016	0.935	1.093	0.000	0.000	0.000	0.000	0.000
3	1.036	1.002	0.978	1.056	1.119	0.000	0.000	0.000	0.000	0.000
4	1.012	1.035	1.017	1.074	1.079	0.000	0.000	0.000	0.000	0.000
Big	1.054	0.882	0.883	0.929	1.159	0.000	0.000	0.000	0.000	0.000
			$\beta_{SMB;i}$					p-value $\beta_{SMB;i}$		
Small	1.666	1.389	1.162	1.061	1.323	0.000	0.000	0.000	0.000	0.000
2	1.264	0.993	0.887	0.878	0.886	0.000	0.000	0.000	0.000	0.000
3	0.828	0.578	0.497	0.589	0.743	0.000	0.000	0.000	0.000	0.000
4	0.373	0.209	0.187	0.337	0.377	0.000	0.000	0.000	0.000	0.000
Big	-0.159	-0.173	-0.127	-0.149	0.170	0.000	0.000	0.010	0.001	0.023
			$\beta_{HML;i}$					p-value $\beta_{HML;i}$		
Small	-0.255	-0.070	0.221	0.544	0.720	0.000	0.215	0.000	0.000	0.000
2	-0.296	0.032	0.279	0.569	0.836	0.000	0.328	0.000	0.000	0.000
3	-0.284	0.102	0.346	0.587	0.916	0.000	0.001	0.000	0.000	0.000
4	-0.271	0.131	0.370	0.519	0.919	0.000	0.000	0.000	0.000	0.000
Big	-0.337	0.033	0.291	0.625	0.869	0.000	0.273	0.000	0.000	0.000
			$\beta_{SOC;i}$					p-value $\beta_{SOC;i}$		
Small	-0.078	0.158	-0.001	0.057	0.535	0.613	0.196	0.994	0.378	0.002
2	-0.167	-0.035	0.080	0.022	-0.073	0.109	0.621	0.226	0.751	0.365
3	-0.274	-0.263	-0.162	-0.195	-0.298	0.002	0.000	0.019	0.003	0.000
4	-0.323	-0.289	-0.326	-0.267	-0.267	0.000	0.000	0.000	0.006	0.005
Big	0.161	-0.028	-0.199	-0.197	-0.067	0.000	0.669	0.008	0.003	0.555
			Adj. R^2							
Small	0.858	0.883	0.931	0.958	0.791					
2	0.920	0.949	0.955	0.950	0.952					
3	0.923	0.950	0.940	0.959	0.955					
4	0.933	0.946	0.936	0.904	0.927					
Big	0.969	0.905	0.891	0.936	0.893					

Table A-6
Governance Risk factor model regression results

$$r_{i;t} = \alpha_i + \beta_{M;i} r_{M;t} + \beta_{SMB;i} r_{SMB;t} + \beta_{HML;i} r_{HML;t} + \beta_{GOV;i} r_{GOV;t} + e_{i;t}$$

	Low	2	3	4	High	Low	2	3	4	High
			α_i (in %)					p-value α_i		
Small	-0.367	-0.036	0.004	0.094	0.517	0.166	0.862	0.980	0.395	0.092
2	-0.117	0.146	-0.022	0.069	-0.035	0.508	0.225	0.843	0.562	0.800
3	-0.193	0.051	-0.103	-0.046	-0.130	0.209	0.654	0.378	0.680	0.365
4	-0.018	-0.009	-0.007	-0.208	-0.095	0.885	0.934	0.957	0.213	0.563
Big	0.216	-0.009	-0.011	-0.160	0.004	0.004	0.937	0.934	0.156	0.982
			$\beta_{M;i}$					p-value $\beta_{M;i}$		
Small	1.023	0.954	0.917	0.876	0.827	0.000	0.000	0.000	0.000	0.000
2	1.088	0.994	1.011	0.934	1.099	0.000	0.000	0.000	0.000	0.000
3	1.056	1.020	0.993	1.070	1.135	0.000	0.000	0.000	0.000	0.000
4	1.036	1.058	1.043	1.092	1.092	0.000	0.000	0.000	0.000	0.000
Big	1.039	0.892	0.896	0.944	1.163	0.000	0.000	0.000	0.000	0.000
			$\beta_{SMB;i}$					p-value $\beta_{SMB;i}$		
Small	1.676	1.371	1.162	1.059	1.273	0.000	0.000	0.000	0.000	0.000
2	1.274	1.000	0.880	0.875	0.891	0.000	0.000	0.000	0.000	0.000
3	0.849	0.599	0.507	0.604	0.772	0.000	0.000	0.000	0.000	0.000
4	0.398	0.229	0.210	0.359	0.404	0.000	0.000	0.000	0.000	0.000
Big	-0.168	-0.178	-0.110	-0.135	0.176	0.000	0.000	0.027	0.002	0.018
			$\beta_{HML;i}$					p-value $\beta_{HML;i}$		
Small	-0.253	-0.074	0.221	0.541	0.702	0.001	0.189	0.000	0.000	0.000
2	-0.290	0.033	0.277	0.569	0.838	0.000	0.319	0.000	0.000	0.000
3	-0.274	0.111	0.352	0.594	0.925	0.000	0.000	0.000	0.000	0.000
4	-0.260	0.141	0.382	0.528	0.927	0.000	0.000	0.000	0.000	0.000
Big	-0.343	0.035	0.298	0.632	0.871	0.000	0.233	0.000	0.000	0.000
			$\beta_{GOV;i}$					p-value $\beta_{GOV;i}$		
Small	0.016	-0.007	-0.001	0.080	0.169	0.913	0.956	0.989	0.209	0.333
2	-0.143	0.056	0.034	-0.007	-0.057	0.160	0.418	0.603	0.923	0.475
3	-0.163	-0.140	-0.144	-0.112	-0.087	0.065	0.032	0.033	0.084	0.287
4	-0.200	-0.206	-0.223	-0.135	-0.049	0.006	0.001	0.002	0.158	0.600
Big	0.153	-0.147	-0.087	-0.125	-0.028	0.000	0.019	0.243	0.053	0.800
			Adj. R^2							
Small	0.858	0.881	0.931	0.958	0.777					
2	0.920	0.949	0.954	0.950	0.952					
3	0.920	0.945	0.939	0.957	0.950					
4	0.927	0.942	0.930	0.900	0.923					
Big	0.969	0.908	0.886	0.934	0.893					