

The effects of Environmental, Social and Governance measures on
the cross section of stock returns, a compensation for risk or
mispricing?

Evidence from the Swedish stock market

Ludvig Annér & Nora Jakobsson van Stam

Supervised by Jian Hua Zhang

Bachelor thesis in finance and economics

Centre of finance at the University of Gothenburg - School of Business,

Economics and Law

Spring term 2018 (15 hp)

Abstract

There is a lack of uniformity throughout the literature regarding the effects of socially responsible investing. By implementing a Fama-MacBeth style regression with the Fama French three factors and the momentum factor, extended with several detailed environmental, social and governance scores the lack of uniformity of these effects are confirmed. The combined social score, the product responsibility and community scores are shown to have positive relations to stock returns, while the human rights and management scores are shown to have negative relations. Deepening the analysis, whether these effects are due to mispricing or risk, there is evidence that the combined social score is to be explained by being a risk factor while the other scores are found to be explained by mispricing.

Contents

1	Introduction	1
1.1	Background	1
1.2	Research Question	2
1.3	Contribution and Purpose	3
2	Literature Review	3
3	Theory	7
3.1	Explanations for the effects of ESG scores on stock returns	7
3.2	CAPM	8
3.3	Carhart's Four-Factor Model	9
3.4	GARCH (1,1) Model	10
3.4.1	ARCH(q) Model	10
3.4.2	GARCH(p,q)	11
4	Methodology	11
4.1	Fama-MacBeth procedure	12
4.2	Identifying Risk-Based Factors	14
4.2.1	The risk factor mimicking portfolios	15
4.3	Robustness	15
5	Data	17
5.1	Stock screening	17
5.2	Carhart's factors	17
5.3	Beta values	17
5.4	ESG data	18
5.5	Risk Mimicking portfolios	19
6	Results	19
6.1	Descriptive Statistics	19
6.2	Effect of ESG scores on stock returns	21
6.3	Risk compensation or mispricing?	25
6.4	Robustness tests	27
7	Conclusions	29
	References	31

A	Grouping Technique	35
B	Autoregressive models	35
B.1	ARCH	35
B.2	GARCH	36
C	The Charoenrook & Conrad methodology	37
D	Thomson Reuters' ESG score	40
E	List of ESG Companies	41
F	Fama-MacBeth in Stata	42
F.1	Beta calculation	42
F.2	The Fama Macbeth regression	42
G	Additional tables and graphs	43

List of Tables

I	Descriptive statistics, Swedish stock market	20
II	Descriptive statistics, ESG sample	20
III	Industry sectors	21
IV	Fama-MacBeth regression, without industry dummies	22
V	Fama-MacBeth regression, with industry dummies	23
VI	Monthly and yearly effects of significant scores	25
VII	Garch estimation	26
VIII	Cross-correlations I	28
IX	Cross-correlations II	28
X	Fama-MacBeth regression, without industry dummies, no grouping technique	43
XI	Fama-MacBeth regression, with industry dummies, no grouping technique . .	44

1. Introduction

1.1. Background

These past years, the awareness of the challenges facing our society and planet, such as climate change, poverty, diseases and increasing income inequalities, has rapidly grown. Governments and societies as whole need to be part of reaching solutions to these issues. Meanwhile, companies and investors play an important role when it comes to distributing capital in the best sustainable way (Generation (2012)). To invest capital sustainably, one strive to maximize the economic value in the long term and the value for shareholders at the same time as one try to preserve and care for environmental and social well-being. To incorporate these external costs in investment decisions, a possibility for investors is to use preferred restrictions and criteria to obtain more sustainable investment. Recent years has shown an increase in the term Environmental, Social and Governance (ESG) criteria, which gives a deeper definition to the Socially Responsible Investing (SRI) concept (MSCI, 2018).

By focusing on sustainable activities, companies can reduce their social and environmental impacts at the same time as it can be used as a tool to improve relations to both employees and investors. A study made by the EY and Boston College Center for Corporate Citizenship (2013) found that firms reporting sustainable activities saw increased company value, they improved in reputation and received better access to capital. The non-financial values have been of more focus recent years and they may help firms to differentiate from competitors. Additionally, previous research has found evidence that by considering sustainable aspects, businesses seem to financially perform above average (Friede, Busch, and Bassen, 2015).

The National Swedish Pension Fund, the AP funds, accounts for about 15 percent of the Swedish national pension system (Fjärde AP-fonden, n.d.). The main task of the AP-funds is to create long-term returns for the pension system. In July 2017, the finance ministry in Sweden created a memorandum, concerning a change in law regarding the first-fourth AP fund. There is a suggestion that, through law, a focus of the first-fourth AP fund should be responsible investments and ownership, with a specific focus on sustainable development. The AP funds shall strive for managing their funds by focusing on how sustainable development can be promoted without compromising on their fiduciary duty of high returns with a

low risk (Finansdepartementet, 2017). As being investment professionals the AP funds has an important role when distributing capital to activities that are of benefit to the society as whole.

Both for companies and investors, professionals and private, it is of importance to gain knowledge in the relation between sustainable activities and stock performances – is it possible to do good and perform financially well at the same time? There is weak consensus in the findings of the studies conducted so far, which suggest that further research is of value. It is also of interest to investigate if the nature of the anomalies is due to a compensation for risk.

1.2. *Research Question*

Several studies have been conducted to measure the effect of sustainable activities on financial performances. However, most of these studies have been conducted in the US regarding the US stock market. Studies and data from the Swedish stock market is not as extensive which makes this market an area of interest. Furthermore, it is of particular interest as Sweden is, according to RobecoSAM (2017), the top country on ESG performance in the world, while the USA is merely ranked on the 14th place. Consequently, this thesis will focus on the Swedish stock market, defined as the companies with Sweden as their country of exchange, and will strive to answer the following research questions:

Do Environmental, Social and Governance scores have an impact on stock returns on the Swedish stock market? In such case, which specific scores and what is the effect?

The first hypothesis tested is:

$H_{0,1}$: Environmental, Social and Governance scores have no effect on stock returns

$H_{a,1}$: Environmental, Social and Governance scores have an effect on stock returns

Previous studies have mainly explained these anomalies qualitatively by risk or mispricing scenarios. As relatively new methods for quantitative deductions have emerged, these can be used to gain additional insights. Therefore, if $H_{0,1}$ can be rejected, possible quantitative explanations for the effects are tested:

Is the explanatory power of the Environmental, Social and Governance measures due to a compensation for risk or mispricing?

The second hypothesis tested is:

$H_{0,2}$: the abnormal return (positive or negative), due to the significant ESG measure, is due to to compensation for risk.

$H_{a,2}$: the abnormal return (positive or negative), due to the significant ESG measure, is due to a mispricing of the market.

1.3. Contribution and Purpose

This study strives to further extend the understanding of ESG measures and its effect on stock returns. By conducting a study on the Swedish stock market as an addition to previous studies, this thesis will contribute with an area where little previous research has been conducted. The updated findings will further contribute to additional insights to this area of continuously increasing focus. Furthermore, this thesis will not merely focus on an aggregated ESG score but also individual, more specific scores are examined. This is of importance for both investors, to see which measures that may have an effect on the returns, and for companies to see which investment areas that may impact shareholder returns.

2. Literature Review

There seem to be no clear consensus of previous conducted studies, investigating the relation between companies' socially responsible (SR) activities and financial performances. There are studies that finds positive relations for companies performing high on SR scores and financial performance, as well as there are studies contradicting these findings, arguing rather the opposite that companies performing poorly on SR scores tend to have higher expected returns.

A recent conducted study by Limkriangkrai, Koh, and Durand (2017) examines the effect of environmental, social, governance and a combined ESG score on stock returns on

the Australian stock market. No effects are found for these aggregated scores. Hence, this supports that further examination on individual score should be conducted.

Manescu (2011) investigates the explanatory power of several ESG measures on stock returns. She further examines whether the significant effects could be explained by mispricing or a compensation for risk. The method used to investigate the effect of ESG measures was a Fama-MacBeth month-by-month, cross-sectional regression. The dependent variable being the monthly stock returns and the independent variables being the four factors suggested by Carhart (1997). The model is further extended with seven ESG measures and controlled for industry sectors. The study finds that the only ESG measure for the full period that shows significance was the community relations, which shows a positive effect on stock returns. Furthermore, the period is separated in two sub periods, 1992-2003 and 2003-2008, which shows positive significant effect for employee relations in the earlier sub period while human rights and product safety shows a negative significant effect in the latter period.

Galema, Plantinga, and Scholtens (2008) also used the Fama-MacBeth procedure to examine the relation between US portfolio returns to different dimensions of SR performance. The study finds a significant positive effect of employee relations on excess returns. Hence, the findings of Manescu (2011) and Galema et al. (2008) support that the underlying characteristics of the social criteria inhibits positive effect for stock returns.

Derwall et al. (2010) conduct a study to investigate the relation between eco-efficiency i.e. the ability to create more value with less environmental resources, and company financial performance, measured as returns on assets, on the US market between 1997 and 2004. The study also examines the change over time. The findings complement the studies above, showing a positive relation between eco-efficiency and operating performance, supporting that environmental aspects also are of significance for asset returns. The result however shows a stronger negative relation between the least eco-efficient companies and operational underperformance than a positive relation for the most eco-efficient companies and positive operational performance, when comparing to a control group. When considering the variation in time, Derwall finds that the more eco-efficient firm the more likely it is to be initially undervalued and then later experience and upward price correction. It is suggested that the time trend is evidence of the market not being able to fully understand the value of SR

activities such as improved eco-efficiency.

Similar result is given for the Polish and Hungarian market (J. Przychodzen and W. Przychodzen, 2014). Companies with high eco-innovation do at most times generate relatively higher returns on assets. Further the study finds that companies that introduce eco-innovation are exposed to lower financial risk and are relatively larger than companies with low eco-innovation.

Edmans (2008), studies the relation between employee satisfaction and long turn stock return in the US from 1984-2009. The author argues based on Markowitz (1959), that screening for SRI would reduce returns as it would restrict the available stock selection. Opposing Markowitz (1959), Edmans also suggests, using human relations theories, that it is sensible to assume a positive relation by screening on SRI criteria. The author interprets human relation theories by claiming that employee satisfaction will increase motivation and retention of the employees. This may increase the efficiency and further the value of the company, which makes Edmans argue that it is rational to assume higher returns for companies performing well on SRI aspects. In consensus with Galema et al. (2008) Edmans finds that firms with high levels of employee satisfaction generate superior returns in the long term. A potential explanation for the positive returns is mispricing, as suggested by Manescu (2011). Higher satisfaction generates higher intrinsic firm value, however the market fails to successfully incorporate this into the stock valuation. Some evidence is found to support that higher return is not merely due to increased satisfaction but the inclusion on SRI lists. Companies included on an SRI list tend to experience higher trading volumes causing increased returns.

The findings of the study made by Kempf and Osthoff (2007) shows similar results. The study compares stocks with high SR ratings to stock with low ratings, by observing the outcomes of going long in the highly rated stocks and going short in the low rated stocks. The high-rated portfolio performs better than the low-rated suggesting investor can earn abnormal returns by using the long-short strategy. The study suggests that the higher returns might be either due to mispricing in the market or a compensation for additional risk.

In contrast, there are studies that suggest that less responsible companies will show greater returns than SR companies. The study of Hong and Kacperczyk (2007) investigate

the effect of social norms on the market by observing so called sin stocks i.e. stocks of companies involved in producing alcohol, tobacco and gambling. The study states that sin stocks could be reasoned to perform better than SR stock. The authors find that sin stocks are less commonly held by institutions constrained by norms e.g. pension funds. The result shows higher expected return for sin stocks compared to similar conventional stocks. The study concludes that norms influence stock returns. Galema et al. (2008) also argues that the higher returns can be explained by the shortage in demand for irresponsible stocks compared to the excess in demand for SR stocks, leading to overpricing of SR stocks and underpricing of sin stocks.

The study of Statman and Glushkov (2009) finds evidence that both SR stocks and sin stocks generates higher returns, when comparing to more conventional investments. The study observes that tilting portfolios towards stocks of high scoring SR companies is of advantage relative to more conventional investors. However, they also observes a disadvantage, relative to conventional investors, when excluding sin stocks. Thus, depending on how the investment is conducted the study shows both negative and positive effects of SR investments.

Both a negative and positive effect of high SR scores on stock returns are seen in the study of Brammer, Brooks, and Pavelin (2006). Using performance indicators for environment, employment and community activities they measures the effect of corporate social performance on stock performance on the Australian stock market. The findings show that firms with higher SR score retrieve lower returns, whereas the firms with the lowest score outperform the market. Considering the individual scores, the environmental and employment measures show a negative relation while community shows a slightly positive relation. The study by Koerniadi, Krishnamurti, and Torani-Rad (2013), conducted in New Zealand, finds that well-governed companies tend to experience lower risk, which could be a reason for the lower returns found by Brammer et al. (2006)

To summarize, previous studies of the relationship of SR investments and firms' financial performance shows various results. The majority of the articles discuss their findings in light of a mispricing or risk compensation scenario, but few quantitative evidence for either scenario is provided. Positive, negative and zero relations are suggested, which makes ex-

pected results of ESG scores hard to predict, albeit contrasting effects of the corresponding sub measures are expected. It seems as if the market is not fully able to coup with ethical measures which makes the area of SR investments an area of interest in need of further exploration.

3. Theory

This theory section is composed as follows. Firstly, in section 3.1 the theories of the possible effects of the ESG scores are discussed. Secondly, the theoretical framework regarding explanatory variables corresponding to the cross section of expected stock returns are presented in section 3.2 and 3.3. Lastly, the underlying theories of the Charoenruek and Conrad (2005) methodology, which is to be used in the identification of risk based factors, is presented in section 3.4.

3.1. Explanations for the effects of ESG scores on stock returns

As suggested by the Literature review, the effect of engaging in SRI is problematic to predict. There is empirical evidence for both zero, negative and positive effects. This section will analyse three possible explanations for the different scenarios.

Firstly, there is the no effect scenario. It suggests that there is no effect of investing in SR stocks on stock returns compared to other stocks. This is in line with the semi-strong efficient market hypothesis, stating that an analysis based on publicly available information should not result in any superior rate of return as the analysis is not likely to be significantly better compared to other analysts' (Bodie, Kane, and Marcus, 2014, pp. 354-356). Therefore, as ESG information has become more publicly available since 2003 (Manescu, 2011), the score should not result in abnormal returns if this information is incorporated efficiently.

Secondly, it is the compensation for risk scenario. The effect of ESG scores on stock returns can be explained by being proxies for risk. Companies rated high on ESG could inhibit either lower or higher risk compared to low rated companies. Engaging in sustainable activities (resulting in increased ESG scores) could increase the risk, by the uncertainty of the net present value of these activities. It could also decrease the risk, by for example,

the companies being better prepared for possible future regulations. If high ESG rated companies carries higher risk, higher expected returns are assumed as a premium for risk. If risk is compensated for in the price, the score is to be considered a risk factor, evidence for this is found in Manescu (2011).

Lastly there is the mispricing scenario to consider, which is the interpreted cause for the high returns related to the environment and employee relations in Derwall et al. (2010) respectively Edmans (2008). Mispricing could cause either higher or lower returns, as market is not able to coup with companies' sustainable activities and consequently it is possible to gain abnormal returns by investing in SR stocks. If the market under- or overestimate the ratio between the cost and benefits of acting sustainably the price on the market will not be efficient. Manescu (2011) argues that if underestimating the benefits while overestimating the costs, companies with high ESG scores will have higher expected return and vice versa.

Considering the effects that are found in previous research, suggested by the literature review, combined with the theories discussed in this section the expected results are as follows. Since ESG information have become more publicly available, the mispricing scenario is the least expected result. The majority of the financial literature on this topic suggest that high performance on sustainable criteria is associated with lower risk¹. Therefore, a negative effect of the ESG scores, corresponding to a lower risk is the expected results. Furthermore, some insignificant results are expected, in line with the no effect scenario.

3.2. CAPM

The Capital Asset Pricing Model (CAPM), created by Sharpe (1964), Lintner (1965) and Mossin (1966), expanding on Markowitz portfolio theory, suggest that the return of an asset must be linearly related to the systematic risk defined as the comovement of the asset returns with the markets. This is estimated by a linear regression with the asset return (r_i) regressed on the market excess return ($RMRF_t$).

$$r_i = \beta_0^i + \beta_1^i RMRF_t + e_i$$

¹See Hong and Kacperczyk (2007), Statman and Glushkov (2009), Brammer et al. (2006), Koerniadi et al. (2013) and Manescu (2011)

The market excess return is proxied by the returns of a large index subtracted with the risk-free rate proxied by the T-bill rate with a relatively short horizon. The coefficient of the excess market return is denoted by 'Beta' in the finance literature.

The Sharpe Ratio is defined as the ratio of excess return of a portfolio ($E[R^p - R^f]$) with its standard deviation ($\sqrt{\text{var}[R^p - R^f]}$) (Sharpe, 1994)²:

$$\text{Sharpe ratio} = \frac{E[R^p - R^f]}{\sqrt{\text{var}[R^p - R^f]}}$$

and is seen as the reward to volatility of a portfolio (Bodie et al., 2014, p. 134).

3.3. Carhart's Four-Factor Model

The CAPM was extended by Fama and French in 1993 with two additional explanatory variables, the market capitalisation and the book-to-market ratio, which provided better explanatory power. Carhart further extended the model by an additional variable, the momentum, proposed by Jegadeesh and Titman (1993) into the four-factor model, resulting in the following model:

$$r_{i,t} = \beta_0^{i,t} + \beta_1^{i,t} \text{RMRF}_t + \beta_2^{i,t} \text{SMB}_{i,t} + \beta_3^{i,t} \text{HML}_{i,t} + \beta_4^{i,t} \text{MOM}_{i,t} + e_{i,t} \quad (\text{i})$$

$r_{i,t}$ = the stock return for firm i in month t.

RMRF_t = the market return in excess in month t.

$\text{SMB}_{i,t}$ = monthly size factor for firm i in month t.

$\text{HML}_{i,t}$ = monthly book-to-market ratio for firm i in month t.

$\text{MOM}_{i,t}$ = monthly momentum factor for firm i in month t.

$e_{i,t}$ = the error term

The *market return in excess* is the value-weighted return of the total market less the risk-free rate. This coefficient is the beta value as estimated in the CAPM. The size and book-to-market variables are constructed with the risk factor mimicking portfolio technique pro-

²This is the ex ante version, the ex post being different in using realized excess returns rather than expected

scribed by Fama and French (1993). To create a portfolio that mimics a risk factor a high minus low procedure is used. It consists of creating several portfolios based on their exposure to a risk factor by then buying the stocks with high exposure to the risk factor and short-selling the stocks with a low exposure to the risk factor.

The *size* risk factor mimicking portfolio is created by splitting the stocks in two portfolios based on the median of the market capitalisation, defined as $Share\ price \times Number\ of\ shares\ outstanding$, and subtracting the average return of the "small" portfolio minus the average return of the "big" portfolio. The factor came from Fama and French's observation that over time small sized firms showed tendencies to outperform large sized firms (Fama and French, 1993). The cross-sectional distribution of a market cap could be problematic to use in a regression analysis, thus the natural logarithm of the market cap is calculated monthly and consequently used (Bali, Engle, and Murray, 2016, p. 89).

The *book-to-market* risk mimicking portfolio is created by splitting the stocks in three based on the 30th and 70th percentiles of their corresponding book-to-market ratio defined as $\frac{Bookvalue}{Marketvalue}$, and constructing a high-minus-low portfolio by taking the high-scoring book-to-market portfolio minus the low-scoring. The ratio is suggested to differentiate between value and growth firms, where a high (low) ratio indicates a value (growth) firm. The variable was created due to the tendency of over performance of value firms relative to growth firms (Fama and French, 1993).

The *momentum factor* refers to the anomaly found by Jegadeesh and Titman (1993). They constructed relative strength portfolios consisting of stock performing well the previous 1-4 quarters and holding them for the consequent 1-4 quarters, resulting in 16 portfolios that were examined. They show that significant abnormal return was present using various portfolio formations using these strategies.

3.4. GARCH (1,1) Model

3.4.1. ARCH(q) Model

The ARCH(q) model was created due to the heteroscedasticity of the error term for some time series. Due to clustering, the error term is dependent on its previous values. Therefore,

the error term can be described as a function of its previous value(s)(Greene, 2003). Since the variance is a function of the errors (ε), the variance (σ_t^2) can be explained by:

$$\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \dots + \alpha_q\varepsilon_{t-q}^2$$

3.4.2. GARCH(p,q)

An extension to the ARCH(q) model is the GARCH(p,q) model. Without going further into the deduction of the model, the explanatory power of the variance can be improved by not only including q lags of the error term (ε), but also including p lags of the variance (σ^2):

$$\sigma_t^2 = \alpha_0 + \delta_1\sigma_{t-1}^2 + \delta_2\sigma_{t-2}^2 + \dots + \delta_p\sigma_{t-p}^2 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\varepsilon_{t-2}^2 + \dots + \alpha_q\varepsilon_{t-q}^2$$

The GARCH(p,q) model has been shown to perform well or better with a small number of terms than an ARCH model with several lags (Greene, 2003, p. 241). A GARCH(1,1) model is specified as:

$$\sigma_t^2 = \alpha_0 + \delta_1\sigma_{t-1}^2 + \alpha_1\varepsilon_{t-1}^2 \quad (\text{ii})$$

For a thorough deduction of the ARCH and GARCH models, see appendix B.

4. Methodology

Similar to Manescu (2011) a cross-sectional regression using the Fama-MacBeth methodology is used. To test *hypothesis 1*, equation (i) is extended with various ESG variables and done in a Fama-MacBeth fashion, which explains the difference between equation (i) and (iii).

$$r_{it} = \beta_0^{it} + \beta_1^{it}Beta_t + \beta_2^{it}SMB_{i,t} + \beta_3^{it}HML_{i,t} + \beta_4^{it}MOM_{i,t} + \beta_5^{it}ESG_{i,t} + e_{i,t} \quad (\text{iii})$$

The hypothesis is rejected if any of the coefficients of the ESG scores are significantly different from zero.

The regression is first performed by observing the effect of the aggregated ESG score as one risk factor. As the ESG score is made up by individual scores the ESG variable is

divided to test for individual effects. Combined environmental, social and governance score are created and tested for, as well as the ten subgroups resource use, emissions, innovation , workforce, human rights, community, product responsibility, management, shareholders and CSR strategy significant explanatory power (further discussed in section 5).

Several empirical studies have shown (Friede et al., 2015) that the importance of ESG on returns may differ across industries. Thus, it is of importance test for industry specific effects. The industry categories are managed in the regression by including the categories as dummy variables. The model is consequently revised to:

$$r_{i,t} = \beta_0^{i,t} + \beta_1^{i,t} Beta_t + \beta_2^{i,t} SMB_{i,t} + \beta_3^{i,t} HML_{i,t} + \beta_4^{i,t} MOM_{i,t} + \beta_5^{i,t} ESG_{i,t} + \sum_{i=1}^{10} \alpha_i^{t+1} Ind_i + e_{i,t} \quad (\text{iv})$$

Where $Ind_i = 1$ if the stock corresponds to the i^{th} sector, and 0 if not.

Based on the data available (see section 5), the tests are additionally split in two sub periods namely, Jan 2003-Dec 2008 and Jan 2009-Dec 2017. This division is chosen for two reasons. Firstly, it is of interest to observe the potential change over time since the awareness of sustainable investment has increased. Secondly, a sample period that excludes the potential disturbing effect of the great recession provides additional value. The data management, regressions and data analysis are performed in the statistical software Stata.

4.1. *Fama-MacBeth procedure*

The models suggested above are estimated using the Fama-MacBeth procedure, which is aimed to estimate the relation between several variables. The Fama-MacBeth regression able us to examine the variable of interest, the ESG scores, while controlling for numerous other variables (Bali et al., 2016, pp. 89-99). One of the main advantages with the Fama-MacBeth procedure suggested by Goyal (2012), is that the analysis can adapt properly to unbalanced panels. In the Fama-MacBeth procedure equal weights for each month are used, thus for an unbalanced data set the model weights all observations proportional to the number of firms for the given month. The regression requires the stocks' beta values which are calculated with the grouping technique proposed by Fama and MacBeth (1973). For a

thorough deduction, see appendix A.

The first step in the two-step procedure is a cross-sectional regression on the dependent variable of interest, the companies' stock returns. The regression will estimate slope coefficients and an intercept coefficient for each regressor for each given period. This results in a time series with slope- and intercept coefficients that are to be saved and used in the next step.

$$r_{1,t} = \beta_0^1 + \beta_1^1 Beta_{1,t} + \beta_2^1 SMB_{1,t} + \beta_3^1 HML_{1,t} + \beta_4^1 MOM_{1,t} + \beta_5^1 ESG_{1,t} + \sum_{i=1}^{10} \alpha_i^{t+1} Ind_i + e_{1,t}$$

$$r_{2,t} = \beta_0^2 + \beta_1^2 Beta_{2,t} + \beta_2^2 SMB_{2,t} + \beta_3^2 HML_{2,t} + \beta_4^2 MOM_{2,t} + \beta_5^2 ESG_{2,t} + \sum_{i=1}^{10} \alpha_i^{t+1} Ind_i + e_{2,t}$$

.

.

.

$$r_{n,t} = \beta_0^n + \beta_1^n Beta_{n,t} + \beta_2^n SMB_{n,t} + \beta_3^n HML_{n,t} + \beta_4^n MOM_{n,t} + \beta_5^n ESG_{n,t} + \sum_{i=1}^{10} \alpha_i^{t+1} Ind_i + e_{n,t} \quad (v)$$

The cross-sectional regression is ran on all companies 1 to N for all time periods 1 to T .

The second step in the analysis is to produce time series averages for the first step's estimated coefficients. The aim is to examine whether the average regression coefficients are statistically different from zero (Bali et al., 2016, p. 91). Any significant difference would indicate a significant relation between the regressor and the dependent variable for the average time-period.

The coefficients that are needed for statistical inference, is the mean values of the $\hat{\beta}_k$ estimates, i.e. $\bar{\beta}_k = \frac{1}{T} \sum_{t=1}^T \beta_t$. The t-statistic is then calculated as $t(\hat{\beta}_k) = \frac{\hat{\beta}_k}{sd(\hat{\beta}_k)/\sqrt{T}}$. For a thorough deduction of this methodology, see Fama and MacBeth (1973). How this is conducted is explained in appendix F.

4.2. Identifying Risk-Based Factors

Charoenruek and Conrad (2005) proposed a model of identifying risk-based factors which is used for testing *hypothesis 2*. They deduct, under certain assumptions, that there must exist a linear relationship between the conditional mean and the conditional variance of the return on a factor-mimicking portfolio, if the factor is a priced risk. By sorting on the proposed risk factor, a risk mimicking portfolio is created by buying (selling) securities with high (low) values of the proposed risk factor.

An assumption of the model is that the portfolio is well diversified, however the authors does not further specify what is considered to be the criterion. Evans and Stephen (1968) argue, based on their findings, that no more than ten securities are needed to create a well diversified portfolio. However, Statman (1987) argues that the portfolio need to consist of at least 30 securities.

The test consists of regressing the portfolio excess returns on it's conditional variance:

$$R_{t+1}^X - R^f = \mu + \delta\sigma_{t+1}^2 + \eta_{t+1} \quad (\text{vi})$$

where $\sigma_{t+1}^2 = \alpha_0 + \delta_1\sigma_t^2 + \alpha_1\varepsilon_t^2$ (equation (ii) in section 3.4)

Based on equation (vi), three criteria are tested for if the ESG score(s) is a risk factor(s), namely:

1. The relation between the conditional mean and variance of the portfolio (captured by δ) should have the same sign as the conditional expected risk premium (estimated as the mean returns) on the risk factor mimicking portfolio. Intuitively, the first criterion is based on two parts: The first one is looking at the relationship between the conditional mean and conditional variance by the δ . Depending on the sign of this relationship and considering that it is constructed in a Low-minus-high fashion (see section 4.2.1), it can be concluded whether the higher risk is associated with higher or lower scores. A positive sign implies a higher risk associated to the lower scores and vice versa.

The second part is that the mean returns of the portfolio should also be of the same sign as the δ . If the δ is positive but the mean negative this suggests that the portfolio is not

compensated risk.

2. The intercept term μ , should not be significantly different from zero, since the expected risk premium for the risk factor mimicking portfolio should be given entirely by its conditional variance, for the deduction of this criteria, see appendix B, by equation C4.

3. The Sharpe ratio should be plausible, where the authors state that the Sharpe ratio should be less than the ex ante tangency portfolio. In a perfect capital market setting, according to MacKinlay (1995), it is sensible for the squared Sharpe measure of the tangency portfolio to be approximately 0.031 for a one-month observation interval. Thus, a plausible squared Sharpe ratio for the factor mimicking portfolio would be less than 0.031.

For a more thorough deduction of their method, see appendix C or Charoenruek and Conrad (2005).

4.2.1. *The risk factor mimicking portfolios*

The risk factor mimicking portfolios are based on the methodology proposed by Fama and French (1993), and further extended by Manescu (2011). The portfolio formation is based on three percentile rankings. The first consist of ranking the stocks based on their relative size, constructing categories Small and Big. Secondly, the same procedure is done with their relative book-to-market values, constructing Growth and Value categories. The last step is ranking based on the ESG risk factor, based on the 30th- and 70th percentile, effectively creating three categories: Low, Medium and High sustainability. This results in twelve ($2 \times 2 \times 3$) portfolios where eight are used to construct the factor mimicking portfolio, namely:

$$\begin{aligned}
 LMH = & \frac{1}{4}(SmallValueLow + SmallGrowthLow + BigValueLow + BigGrowthLow) \\
 & - \frac{1}{4}(SmallValueHigh + SmallGrowthHigh + BigValueHigh + BigGrowthHigh)
 \end{aligned}
 \tag{vii}$$

4.3. *Robustness*

To check the quality of the method and accuracy of the estimates a few robustness tests are conducted. Petersen (2009) shows that the Fama-MacBeth procedure creates a downward

bias of the standard errors in the case of unobserved firm specific effects, i.e. that the errors of the firms are correlated over time. In the case of unobserved time effects, i.e. that the errors are correlated across firms for a given time, the procedure creates no bias. Petersen provides an example of a data set with similar characteristics as the data set of this thesis, where it is shown that these characteristics provide unbiased standard errors in the case of a Fama-MacBeth regression as no unobserved firm specific effect is present. Thus, the potential bias in the standard errors should be avoided. Bali et al. (2016, p. 91) suggests using Newey-West standard errors to avoid problems with potential heteroscedasticity in the error term which can cause incorrect standard errors, therefore these are used for all regressions with a lag of 12.

The Fama-MacBeth procedure assumes, at most times, a linear relation between the regressors and the outcome, in such case a OLS regression should be used for the cross-sectional analysis (Bali et al., 2016, p. 89). Thus, a test for linearity is conducted. The regressors are tested for multicollinearity, when two or more of the regressors are correlated to one another. If any correlation is strong, the overall accuracy of the model is not affected, however the estimated coefficients of the correlated variables may be inaccurate.

When using the grouping technique to estimate the beta values, a limited number of portfolios are formed. Due to this potential lack of precision in the portfolio betas, the Fama-MacBeth regression is also performed without the grouping technique.

As mentioned in the Literature review, inclusion on an SRI list could be the factor effecting expected stock returns. Henceforth, all companies on the Swedish stock market are used in a regression to see if having an ESG score is the reason for difference in stock returns. This is tested by a Fama MacBeth regression by including an ESG dummy variable to Carhart's four factors.

5. Data

5.1. Stock screening

The individual firm stock returns are gathered monthly with the total return function from Thomson Reuters Eikon (2018). All firms are retrieved from the Swedish stock market, by sorting on country of exchange. Stocks are screened for ESG measures, if having an ESG Score they are included, creating an unbalanced panel. After the screening 69 firms remained for the period Jan 2003-Dec 2017, this is the period for when the ESG scores of Thomson Reuters are available for the Swedish stock market. If the monthly returns are unavailable, the observations are dropped resulting in 779 dropped observations. In total, 11,641 firm-month observations are gathered.

5.2. Carhart's factors

All data for the factors are gathered from Thomson Reuters Eikon on a monthly basis. To receive excess market returns the monthly risk-free rate is subtracted from the market returns. 3-month Treasury Bills are gathered from Sveriges Riksbank (2018) as a proxy for the risk-free rate, and made monthly by the following formula:

$$\text{Monthly rate} = ((1 + \text{Annual rate})^{1/12}) - 1$$

The Momentum variable is calculated by the average of the one month total return for month $t - 12$ to $t - 2$, $\frac{1}{|t|-1} \sum_{t=-12}^{t=-2} R_t$. Beta values are calculated with the grouping technique as suggested in the section 4.1 and elaborated in appendix A.

5.3. Beta values

The first step in the grouping technique, is the portfolio creation based on firm specific beta values. Monthly regressions are made for each asset. However, since the beta is highly sensitive in the beginning of an assets life due to a small number of data points, only betas that are calculated with at least 36 data points (i.e. 3 years of monthly data) are kept.

The number of portfolios formed are modified slightly from Fama and MacBeth (1973) and Manescu (2011) (20 and 50 portfolios respectively) since, as further explained in appendix A, the methodology is dependent on portfolios with sufficient number of securities to essentially remove the variance of the error term, it would be of small value to create 20 portfolios and of no value to create 50, due to the small number of securities that would be corresponding to each portfolio. Therefore, ten portfolios are created based on their individual betas.

5.4. *ESG data*

To ensure relevant and transparent ESG data, Thomson Reuters' ESG scores are used. The data is gathered through Thomson Reuters Eikon, where ESG scores have been conducted since 2002. The aim with the score is to structure and create standardised measures for ESG data which can be used for financial analysis. Most of the data used comes from publicly reported information such as annual and CSR reports. The score also includes exclusion criteria, such as alcohol, armaments and gambling. Using more than 150 research analysts Thomson Reuters states they obtain one of the largest ESG content collection operations in the world (Thomson Reuters, 2018). The ESG scores measures performances in ten main categories, stated in section 4, made up by 178 comparable measures. These ten categories are combined into an environmental, a social and a governance score. An aggregated ESG score is also provided, made up by all ten individual categories. The categories have different weights in the combined scores (see appendix D).

The process of creating an ESG score has resulted in three numerical values for all the screened firms. Firstly, there is the score, which provides a numerical value between 0 to 100 for each category, with the higher score indicating the better performance. The data making up the score is initially derived from the firms' financial reports. Secondly, there is the percentile rank. Based on the ten subgroups, percentile ranks are calculated for all the screened firms. Finally, there is the ratings, which is a relative ranking depending on the other companies. The firm rating can be used to compare ESG measures to other firms, it can also be used for specific category comparison to obtain a proper measure of a firm's environmental, social, governance or combined ESG activities. The ESG scores are updated

yearly in contrast to the other variables. For a more thorough deduction of Thomson Reuters ESG data and methodology, see Blank (2013) and Thomson Reuters (2018).

Industry sectors, used as control variables, are gathered from Eikon and based on the Global Industry Classification Standard (GICS).

5.5. *Risk Mimicking portfolios*

The majority of the risk factor mimicking portfolios are created as proposed in section 4.2.1, ie. based on the 30th-, and 70th percentiles. However, due to limitations in the GARCH estimation procedure, the product responsibility (whole period) and the human rights (latter subperiod) risk mimicking portfolio was formed based on the 20th-, and 80th percentiles. The management risk mimicking portfolio (latter subperiod) was formed based on the 5th-, and 95th percentile. The portfolio constructed based on the 30th- and 70th percentile results in portfolios consisting of 41 stocks, the portfolios based on the 20th- and 80th percentile in 27 stocks, and finally the portfolios based on the 5th- and 95th percentile in only 7 stocks.

6. Results

6.1. *Descriptive Statistics*

Some descriptive statistics for all companies on the Swedish stock market and for the screened ESG companies are provided by table I and II respectively. After inspection of the variables and following the guidance of Bali et al. (2016, p. 90), the right tail of book-to-market and both tails of momentum at the 0.5% percentile are winsorized due to them having clear outliers. For the robustness tests with all companies on the Swedish stock market the dependent variable (monthly returns) is also winsorized due to large outliers in this variable which otherwise could cause errors in variables bias as discussed by Bailer and Martin (2007).

The mean of the size factor (23.518) for the companies with an ESG score indicates that on average companies with an ESG score are large cap companies while the average companies on the Swedish stock market are small cap companies (Nasdaq, 2017).

Population: 755	mean	σ	min	max	p1	p25	p50	p75	p99
Total return	2.148	105.814	-97.960	24,900	-35.940	-6.383	-0.000	6.731	67.033
Rm-Rf	0.646	4.435	-18.183	18.705	-12.825	-1.570	1.041	3.134	10.122
Size	20.276	2.418	11.821	26.999	15.620	18.497	19.994	21.922	26.130
Book-to-market	0.334	36.947	-3,366.913	1,768.692	-0.032	0.230	0.458	0.855	8.810
Momentum	2.267	33.841	-25.137	2,342.584	-11.621	-1.279	1.343	3.853	21.308

Table I: Mean, standard deviation (σ), minimum value, maximum value, 25th, 50th, 75th percentile values for the asset returns, market excess return, size (natural log of the market capitalisation), book-to-market and the momentum for the time period Jan 2003-Dec 2017.

Sample size: 69	mean	σ	min	max	p25	p50	p75
Monthly return	1.562	10.204	-52.469	195.082	-3.636	1.159	6.280
Rm-Rf	0.714	4.630	-18.183	18.705	-1.699	0.988	3.434
Size	23.546	1.680	17.259	26.999	22.542	23.669	24.660
Book-to-market	0.687	0.805	-0.473	16.686	0.288	0.515	0.816
Momentum	1.536	3.520	-21.640	32.209	-0.191	1.600	3.201
ESG	58.376	14.959	8.108	86.727	48.604	60.421	69.758

Table II: Mean, standard deviation (σ), minimum value, maximum value, 25th, 50th, 75th percentile values for the asset returns, market excess return, size (natural log of the market capitalisation), book-to-market, momentum and the aggregated ESG score for the time period Jan 2003-Dec 2017.

In the study by Banz (1981) it is found that smaller companies tend to have on average higher risk adjusted returns compared to large companies. No clear difference was seen between medium sized and large companies. This differences in average company size may be a reason for difference in expected returns of the ESG sample and the whole market, supported by the mean returns in table I and II.

Considering the industry sectors, shown in table III, for the companies with an ESG score, the industrial sector is clearly the largest sector, making up about 35 percent of the observations. For the entire Swedish stock market, health care and information technology makes up 42.8 percent of the observations, while they only make up 11.6 percent of the observations for the ESG sample. We hypothesize that this difference is mainly due to the tendency of companies with an ESG score to be larger than for the full sample which is evident from the size means as described above. This is supported when calculating the means of all companies on the Swedish stock market, as it becomes evident that information technology companies have the smallest mean (19.34), health care takes the second to

Sector	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	Total
All	98 (13.01)	23 (3.05)	17 (2.26)	34 (4.52)	171 (22.71)	154 (20.45)	149 (19.79)	31 (4.12)	62 (8.23)	10 (1.33)	4 (0.53)	755 (100.00)
ESG	9 (13.04)	4 (5.80)	1 (1.45)	9 (13.04)	5 (7.25)	23 (33.33)	3 (4.35)	6 (8.70)	7 (10.14)	2 (2.90)	0 (0.00)	69 (100.00)

Table III: Industry sector frequencies the ESG sample and all companies on the Swedish stock exchange in Dec 2017. Based on Global Industry Classification Standards. Percentages in parenthesis. Industry sectors: 1. Consumer discretionary 2. Consumer staples 3. Energy 4. Financials 5. Health care 6. Industrials 7. Information technology 8. Materials 9. Real estate 10. Telecommunication 11. Utilities

last position (19.41) while the industrials takes the fifth position (20.54). The descriptive statistics suggests a sample selection bias considering the differences between the means and returns in table I and II, and the different sector belongings in table III. As the effects of these variables are controlled for in the size and industry dummy variables in the regression, this is not an issue.

6.2. Effect of ESG scores on stock returns

Table IV and V shows the results from the regressions, testing for the first hypothesis, when not controlling and when controlling for industry sectors respectively. Marginal statistical significance, between five and ten percent level, is found in some scores, however these has not been further examined as it is consider by many that no conclusions should be drawn at this level of significance (Bali et al., 2016, p. 96). The results will be discussed variable-by-variable.

For the compounded scores environmental, social and governance, only marginal effects are found for the scores social and governance, when observing the whole period, but it diminishes when controlling for industry sectors. When observing the individual ESG scores for the whole period, the management score is the only score with significant explanatory power. However, the effect disappears when controlling for the industry sectors. This indicates that these effects are partially due to industry characteristics of the companies with these scores. The lack of signifiance for the ESG variables for the whole period are in line with the no effect scenario, i.e. the market has efficiently incorporated the ESG scores in the stock pricing.

Without industry dummies									
Variables	Jan 2003-Dec 2017			Jan 2003-Dec 2008			Jan 2009-Dec 2017		
Beta	-0.013 (0.970)	-0.004 (0.990)	-0.028 (0.941)	-0.198 (0.794)	-0.056 (0.944)	-0.262 (0.750)	0.143 (0.632)	0.095 (0.742)	0.258 (0.531)
Size	-0.178 (0.206)	-0.228 (0.129)	-0.291* (0.086)	-0.243 (0.361)	-0.321 (0.282)	-0.509 (0.107)	-0.144 (0.345)	-0.173 (0.249)	-0.137 (0.425)
Book-to-market	0.002 (0.994)	0.003 (0.987)	-0.052 (0.819)	-0.132 (0.306)	-0.139 (0.258)	-0.153 (0.484)	0.177 (0.617)	0.172 (0.621)	0.083 (0.823)
Momentum	0.122 (0.104)	0.115 (0.141)	0.157** (0.034)	0.043 (0.596)	0.039 (0.665)	0.109 (0.378)	0.162 (0.157)	0.159 (0.170)	0.183* (0.054)
ESG (1-10)	0.000 (0.952)			0.001 (0.934)			0.004 (0.614)		
Environmental (1-3)		-0.002 (0.720)			0.003 (0.771)			-0.006 (0.425)	
Social (4-7)		0.010* (0.065)			0.010 (0.347)			0.012* (0.068)	
Governance (8-10)		-0.008** (0.050)			-0.013* (0.062)			-0.001 (0.726)	
1. Resource use			0.004 (0.371)			0.005 (0.502)			0.002 (0.668)
2. Emissions			-0.004 (0.472)			-0.005 (0.678)			-0.006 (0.185)
3. Innovation			-0.001 (0.903)			0.006 (0.415)			-0.003 (0.786)
4. Workforce			0.001 (0.884)			-0.005 (0.654)			0.008 (0.207)
5. Human rights			-0.006 (0.380)			0.013* (0.067)			-0.018*** (0.007)
6. Community			0.006* (0.092)			0.005 (0.369)			0.008** (0.050)
7. Product responsibility			0.004 (0.209)			0.011** (0.046)			-0.002 (0.522)
8. Management			-0.010** (0.017)			-0.016** (0.043)			-0.004 (0.149)
9. Shareholders			0.003 (0.230)			0.006 (0.166)			0.003 (0.429)
10. CSR strategy			0.006 (0.140)			0.001 (0.882)			0.008 (0.102)
Constant	5.258 (0.106)	6.429* (0.065)	8.091** (0.030)	7.049 (0.267)	8.789 (0.218)	12.397* (0.094)	3.965 (0.243)	4.701 (0.158)	4.577 (0.194)
Observations	7,809	7,809	7,809	2,378	2,378	2,378	5,430	5,430	5,430
Groups	180	180	180	72	72	72	108	108	108
Average R ²	0.242	0.287	0.448	0.258	0.314	0.526	0.233	0.270	0.395

Table IV: Equation (iv) estimated with the Fama-Macbeth Procedure for the whole time period and the sub periods, without controlling for industry sectors. The returns (dependent variable) are in percentage form. P-values in parenthesis, *** p<0.01, ** p<0.05, * p<0.1. Beta values estimated with the grouping technique with 10 portfolios. Groups is the number of time intervals (months).

With industry dummies									
Variables	Jan 2003-Dec 2017			Jan 2003-Dec 2008			Jan 2009-Dec 2017		
Beta	-0.253 (0.488)	-0.203 (0.589)	-1.526 (0.293)	-0.533 (0.459)	-0.478 (0.487)	-3.446 (0.322)	-0.108 (0.749)	0.016 (0.964)	0.081 (0.798)
Size	0.027 (0.883)	0.011 (0.949)	0.210 (0.436)	0.086 (0.788)	0.074 (0.775)	0.311 (0.576)	-0.084 (0.733)	-0.110 (0.670)	0.071 (0.792)
Book-to-market	-0.082 (0.797)	-0.114 (0.723)	-0.368 (0.414)	-0.356 (0.364)	-0.417 (0.332)	-0.540 (0.563)	0.012 (0.977)	0.001 (0.997)	-0.078 (0.851)
Momentum	0.137 (0.163)	0.137 (0.145)	-0.000 (0.998)	0.058 (0.569)	0.053 (0.524)	-0.247 (0.480)	0.181 (0.231)	0.195 (0.186)	0.205 (0.111)
ESG (1-10)	-0.008 (0.368)			-0.011 (0.441)			0.001 (0.968)		
Environmental (1-3)	-0.004 (0.616)			0.008 (0.523)			-0.013 (0.204)		
Social (4-7)	0.001 (0.803)			-0.010 (0.155)			0.015** (0.028)		
Governance (8-10)	-0.005 (0.372)			-0.010 (0.345)			0.000 (0.924)		
1. Resource use	-0.003 (0.844)			-0.016 (0.589)			0.001 (0.881)		
2. Emissions	-0.014 (0.214)			-0.029 (0.282)			-0.004 (0.288)		
3. Innovation	-0.003 (0.716)			0.012 (0.402)			-0.010 (0.315)		
4. Workforce	0.009 (0.547)			0.018 (0.589)			0.005 (0.441)		
5. Human rights	-0.013 (0.252)			0.004 (0.867)			-0.016*** (0.002)		
6. Community	-0.004 (0.652)			-0.023 (0.170)			0.011*** (0.007)		
7. Product responsibility	0.007 (0.268)			0.025** (0.030)			-0.004* (0.092)		
8. Management	-0.006 (0.492)			-0.004 (0.857)			-0.005** (0.011)		
9. Shareholders	-0.003 (0.751)			-0.011 (0.660)			0.004 (0.364)		
10. CSR strategy	-0.001 (0.905)			-0.013 (0.425)			0.006 (0.216)		
Constant	0.998 (0.805)	1.411 (0.704)	4.312 (0.460)	-0.781 (0.915)	0.837 (0.879)	12.299 (0.274)	2.898 (0.580)	3.149 (0.574)	-0.787 (0.895)
Observations	7,809	7,809	7,809	2,378	2,378	2,378	5,430	5,430	5,430
Groups	180	180	180	72	72	72	108	108	108
Average R ²	0.533	0.567	0.692	0.626	0.668	0.816	0.470	0.497	0.608

Table V: Equation (iv) estimated with the Fama-Macbeth Procedure for the whole time period and the sub periods, controlling for industry sectors. The returns (dependent variable) are in percentage form. P-values in parenthesis, *** p<0.01, ** p<0.05, * p<0.1. Beta values estimated with the grouping technique with 10 portfolios. Groups is the number of time intervals (months).

Since the findings suggest that there exist industry effects, the regressions controlling for industry sector are considered to be most reliable. Based on said regressions, there are five significant scores. Both positive and negative effects are found (three and two respectively) which strengthens the lack of uniformity of the effects of ESG measures found in previous studies.

The aggregated ESG score does not have any significant impact on stock returns since it is insignificant in the whole period and both sub periods which is in line with the findings of Manescu (2011) and Limkriangkrai et al. (2017). Furthermore, considering the different signs of the individual ESG scores, this becomes rather intuitive. Observing the sub periods, the social score has significant positive effect for the latter period. Thus, 33 percent of the combined environmental, social and governance scores show to have an effect on stock returns. The environmental and governance performance of the companies are therefore assumed to be fully incorporated into the stock prices. Regarding the individual scores, product responsibility is the only score with significance in the first period, with a positive coefficient. For the second period human rights and management are negatively related with the stock returns, this is in line with Manescu (2011) and Koerniadi et al. (2013) respectively. The community score is showed to have a significantly positive effect in the second period, the same effect is found in both Manescu (2011) and Brammer et al. (2006). Hence, 40 percent of the individual scores have an impact on stock returns. The other individual scores, that show no significant effects, should therefore be fully incorporated into the stock prices, in accordance with the no effects scenario.

It is important to note the differences between the time periods. Due to the greater significance in the latter period compared to the whole and the first, an increase in the effects of ESG scores is observed. It may be due to investors and companies being increasingly engaged in SR investments and activities. Another possible explanation, perhaps in conjunction with the first, is that there are both more companies with an ESG score, and that the extent of the reporting have increased. This makes the ESG score increasingly important for the determination of stock returns. However, previous studies find significant effects of SR scores for earlier time periods which suggest that SR activities have been effecting stock returns for a longer period than these results suggest.

Score	Jan 2003-Dec 2017				Jan 2003-Dec 2008				Jan 2009-Dec 2017			
	SD	$\bar{\beta}$	Month	Year	SD	$\bar{\beta}$	Month	Year	SD	$\bar{\beta}$	Month	Year
Social	19.875	-	-	-	17.676	-	-	-	20.651	0.015	0.310	3.717
Human rights	23.978	-	-	-	22.497	-	-	-	22.423	-0.016	-0.359	-4.305
Community	28.214	-	-	-	28.927	-	-	-	27.614	0.011	0.304	3.645
Product respons.	27.216	-	-	-	24.332	0.025	0.608	7.300	28.440	-	-	-
Management	28.462	-	-	-	28.852	-	-	-	28.281	-0.005	-0.141	-1.697
Groups	180	180	180	180	72	72	72	72	108	108	108	108

Table VI: Standard deviation (SD), estimated coefficient ($\bar{\beta}$), monthly and yearly effects of a one-standard-deviation change in the significant ESG scores in table IV and V. Groups is the number of time intervals (months).

The estimated marginal effect on stock returns of a one-standard-deviation increase in the found significant scores (p-value<0.05) are provided in table (VI), both on a monthly and yearly basis. Comparing these results to previous studies discussed, these ESG effects are rather extreme, especially product responsibility for the first period (7.300) and human rights for the latter period (-4.305). The effects of the community and social for the latter period are also quite high (3.717 and 3.645 respectively). A potential explanation for the difference in the magnitude of the effects could be differences when conducting the studies, as mentioned above. Furthermore, there is also a possibility that the Swedish market values sustainable activities in a different fashion. Since no previous studies on this topic conducted on the Swedish market have been found the latter explanation cannot be disregarded.

6.3. Risk compensation or mispricing?

Using the procedure proposed by Charoenrook and Conrad (2005), the second hypothesis is tested, i.e. the significant effects of the ESG scores are examined for either being due a compensation for risk or mispricing. The model requires the assumption of a well diversified portfolio, as described in section 4.2. The portfolios based on the 30th- and 70th percentile is well diversified (41 stocks), the 20th- and 80th percentile is on the lower spectrum (27 stocks). The 5th- and 95th percentile is not well diversified (7 stocks) and should therefore be disregarded, this is indicated by strikeouts in table VII. A graphical representation of the regressions with significant δ values is provided in appendix G.

The necessary data to distinguish if the effects of the ESG scores are due to risk or

Score	Jan 2003-Dec 2017				Jan 2003-Dec 2008				Jan 2009-Dec 2017			
	μ	δ	Mean Sharpe ²		μ	δ	Mean Sharpe ²		μ	δ	Mean Sharpe ²	
Social	.004 (.994)	-.044 (.173)	-.576	.026	-2.201* (.080)	.118 (.123)	-.404	.011	.639 (.394)	-.108** (.049)	-.680	.035
Human rights	1.278 (.113)	-.009** (.047)	-.211	.003	2.971 (.463)	-.143 (.379)	-.555	.012	-.850 (.577)	.138 (.192)	1.087	.082
Community	-.864 (.214)	.037 (.359)	-.287	.005	-3.328 (.781)	.121 (.812)	-.48	.010	1.047 (.538)	-.172 (.336)	-.564	.031
Product respons.	-1.129 (.699)	.036 (.807)	-.423	.009	-1.266 (.783)	.023 (.9)	-.693	.020	2.265 (.200)	-.163 (.177)	-.072	.000
Management	-3.033** (.025)	.193*** (.009)	.405	.090	.251 (.852)	-.011 (.858)	.029	.000	.391 (.844)	-.015 (.828)	-.027	.000
Groups	180	180	180	180	72	72	72	72	108	108	108	108

Table VII: The intercept μ and coefficient δ from a simple linear regression of equation (vi), mean values of the factor mimicking portfolios based on equation (vii) and the squared Sharpe ratio, for the periods Jan 2003-Dec 2017, Jan 2003-Dec 2008 and Jan 2009-Dec 2017. Groups is the number of time intervals (months). P-values in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

mispricing according to the criteria given in section 4.2 is provided in table VII. Furthermore, the results are compared to the findings in the Fama-MacBeth regression to examine whether the effects of the two tests corresponds, in line with the analysis in Manescu (2011).

Considering the *social score*, in the second period, the first criterion for being a risk factor is fulfilled since the δ (-0.108) is significant and the sign of the coefficient is the same as the sign of the mean (-0.680). The negative sign of the δ for the low-minus-high portfolio implies that the risk is associated with higher social scores. As the μ is insignificant the second criterion is also fulfilled implying that the effect is entirely given by the conditional variance. For the third criterion the squared Sharpe value (0.035) exceeds the suggested upper bound (0.031). Considering the marginal difference, the Sharpe value is considered plausible and the third criterion is therefore fulfilled. Hence, there is strong evidence that the social score is a priced risk factor. Furthermore, the negative sign of the δ (a higher score indicates higher risk) corresponds with the positive sign of the coefficient of the social score from the Fama-MacBeth regression (indicating that the return increases with higher social score) in table V.

For the *human rights score*, for the whole period, all three risk factor criteria are fulfilled. The δ (-0.009) is significant and its negative sign corresponds with the negative sign of the mean (-0.211). The μ is insignificant and the squared Sharpe value (0.003) is considered

plausible, albeit low. The negative δ , indicates that high risk is associated with high human rights score. Companies with high human rights scores are therefore expected to generate higher returns than companies with low scores as a compensation for risk. This does not however, correspond with the negative effect from the Fama-MacBeth regression. Hence, the human rights score cannot be considered a risk factor.

Neither the *community* nor the *product responsibility score* fulfil any criteria for being a risk factor. Thus, the effects of both these scores are argued being caused by mispricing.

The *management score*, for the whole period, fulfils the first criterion, the δ (0.193) is significant and its positive sign corresponds with the positive mean value (0.405). However, the μ (-3.033) is significantly different from zero i.e. the effect of the management score is not entirely given by the conditional variance, indicating that there are others factors than the risk driving the returns. The squared Sharpe value (0.090) greater than the proposed upper bound (0.031) by almost a factor of three and is therefore not considered plausible. Thus, there is only weak evidence for the management score being a risk factor.

Hence, one out of five (20 percent) of the scores found with an significant effect on stock returns are considered being a priced risk factor. The effects of the other scores are argued being due to mispricing.

6.4. *Robustness tests*

Linearity between the dependent variable and the independent variables were tested by plotting the different independent variables to the monthly returns, the dependent variable. Approximate linearity is found for all relations and therefore the underlying regressions are done in an OLS fashion.

Variables	Monthly return	Beta	Size	Book-to-market	Momentum	ESG
Monthly return	1.000					
Beta	0.479	1.000				
Size	0.006	0.007	1.000			
Book-to-market	-0.086	-0.040	-0.084	1.000		
Momentum	0.030	0.028	0.046	-0.254	1.000	
ESG	-0.009	0.002	0.398	-0.019	-0.069	1.000

Table VIII: Cross-correlations between the monthly returns, beta, size, book-to-market, momentum, and ESG, for the period Jan 2003-Dec 2017.

Variables	ESG	Resource use	Emissions	Innovation	Workforce	Human rights	Community	Product responsibility	Management	Shareholders	CSR strategy
ESG	1.000										
Resource	0.738	1.000									
Emissions	0.679	0.589	1.000								
Innovation	0.515	0.417	0.408	1.000							
Workforce	0.709	0.535	0.444	0.262	1.000						
Human ri.	0.543	0.567	0.426	0.305	0.306	1.000					
Community	0.572	0.424	0.294	0.191	0.345	0.425	1.000				
Product re.	0.507	0.300	0.314	0.211	0.387	0.243	0.294	1.000			
Management	0.460	0.038	0.072	0.041	0.039	0.022	0.128	0.092	1.000		
Shareholders	0.209	0.066	0.002	-0.006	0.087	0.013	0.014	-0.063	0.119	1.000	
CSR strat.	0.598	0.495	0.361	0.216	0.482	0.419	0.421	0.322	0.122	-0.031	1.000

Table IX: The cross-correlations between the variables ESG, resource use, emissions, environmental innovation, workforce, human rights, community, product responsibility, management, shareholders and CSR strategy scores, for the period Jan 2003-Dec 2017.

Testing for *multicollinearity* by observing the cross-correlations of the independent variables, only relatively weak relations were found, see table VIII and IX. Thus, multicollinearity should not be an issue effecting the estimations.

Edmans (2008) argues that the effect of the ESG score could be the effect of being included in an ESG/SR-list which attracts investment and therefore effects returns, and not merely the score in itself. By running the Fama MacBeth regression described in 4.3 on all the companies on the Swedish stock market (excluding non-sector belonging companies) this potential effect is tested for. The results show a positive marginally significant effect (p-value=0.087) on the dummy variable, however, as the sample was limited based on size, this effect diminishes³. This suggests that the effect proposed by Edmans (2008) is not present

³The results can be provided upon request

in the Swedish market for the period Jan 2003-Dec 2017.

Due to the limited number portfolios that were formed in the grouping technique, to remove the error in variables bias, the Fama-MacBeth regressions are also computed without the grouping technique, i.e. using their beta values that contains the error term with potential attenuation bias. The results are shown in table X and XI in appendix G. All ESG variables are consistent in both regressions excluding product responsibility that is not significant for the earlier sub period. The effect of product responsibility should hence be taken with caution.

7. Conclusions

This thesis is investigating the effects of environmental, social and governance (ESG) scores on stock returns, and whether the effects could be explained by risk or mispricing. The focus is the Swedish stock market for the period Jan 2003-Dec 2017. The sample is further split into two sub periods, Jan 2003-Dec 2008 and Jan 2009-Dec 2017.

To test for the effects of the ESG scores on stock returns, a Fama-MacBeth regression with Carhart's four factors, extended with ESG scores and industry dummies to control for industry specific effects is performed. Firstly, an aggregated score made up by the ten individual scores⁴, provided by Thomson Reuters Eikon (2018), is used in the regression. Secondly, combined environmental, social and governance scores are created with the ten individual scores⁵ and are used. Lastly, the ten individual ESG scores provided by Thomson Reuters Eikon (2018) are used in the regression.

The aggregated ESG score do not have significant effect for any time period. For the combined environmental, social and governance score, only the social score (33 percent) shows to have a significant effect on stock returns for the second period. Furthermore, significant positive or negative effects on stock returns are found for 40 percent of the individual scores, namely the human rights, community, product responsibility and management score. Prod-

⁴The individual scores are: 1. Resource use, 2. Emissions, 3. Innovation , 4. Workforce, 5. Human rights, 6. Community, 7. Product responsibility, 8. Management, 9.Shareholders and 9. CSR strategy

⁵Environmental is created by score 1-3, social is created with score 4-7 and governance is created with score 8-10.

uct responsibility is the only score with a significant effect on stock returns in the first sub period while the human rights, community and management scores have a significant effect on stock returns in the second period. These findings are in line with previous research which found confounding effects between various ESG scores and stock returns. Furthermore, due to the confounding significant effects of the individual ESG scores and the lack of significance in the aggregated ESG score, it is concluded that in order to better examine the effects of ESG performance on stock returns, more specific scores than one aggregated should be used.

Testing for whether the effects are due to risk or mispricing, the procedure of Charoenruek and Conrad (2005) is used. There is found evidence that the effects of the individual scores, the human rights, community, product responsibility and management score are caused by mispricing. Considering the combined social score, evidence for compensation of risk is found. These evidence for risk compensation and the findings that the majority of the ESG variables (64 percent) have no impact on stock returns suggest that the market efficiently incorporates ESG information in its price setting. However, as there is evidence that the effects of some of the individual scores, are due to mispricing of the market, it is concluded that the market for the second sub period is not fully efficient in its incorporation of ESG related information.

This thesis provides insights of the effects and their nature of several detailed ESG scores. These insights can be used by investors in their decision of which weight should be given to the assets' sustainability performance. By conducting the study and finding significant results on the previously relatively unresearched Swedish market, this thesis broadens the understanding of sustainable investments.

A limitation in this study is the relatively limited sample size. This is not surprising as the focus is limited on the Swedish stock market. To avoid this issue and still investigate a market where sustainability performance is high, further research could focus on the Scandinavian or Nordic market.

References

- Bailer, H. and Martin, R. D. (2007) Fama macbeth 1973: reproduction, extension, robustification, *Journal of Economic and Social Measurement*, 41–63.
- Bali, T. G., Engle, R. F., and Murray, S. (2016) Empirical asset pricing – the cross section of stock returns. New Jersey: John Wiley Sons, Inc.
- Banz, R. W. (1981) The relationship between return and market value of common stocks, *Journal of Financial Economics* **9**, 3–18.
- Black, F., Michael, J. C., and Scholes, M. (1972) The capital asset pricing model: some empirical tests, *Praeger Publishers Inc.*
- Blank, H. (2013) Thomson reuters corporate responsibility ratings. URL: <https://financial.thomsonreuters.com/content/dam/openweb/documents/pdf/tr-com-financial/methodology/corporate-responsibility-ratings.pdf>. (accessed: 2018-03-15).
- Bodie, Z., Kane, A., and Marcus, A. J. (2014) Investments. New York: McGraw-Hill Education.
- Brammer, S., Brooks, C., and Pavelin, S. (2006) Corporate social performance and stock returns: uk evidence from disaggregate measures, *Financial management* **35**, 97–116.
- Carhart, M. M. (1997) On persistence in mutual fund performances, *The Journal of Finance* **52**, 57–82.
- Charoenrook, A. and Conrad, J. S. (2005) Identifying risk-based factors, *SSRN Electronic Journal*.
- Derwall, J., Guenster, N., Bauer, R., and Koedijk, K. (2010) The economic value of corporate eco-efficiency, *European Financial Management* **17**, 679–704.
- Edmans, A. (2008) Does the stock market fully value intangibles? employee satisfaction and equity prices, *Journal of Financial Economics* **101**, 621–640.
- Evans, J. L. and Stephen, H. A. (1968) Diversification and the reduction of dispersion: an empirical analysis, *The Journal of Finance* **23**, 761–767.
- EY and Boston College Center for Corporate Citizenship, B. c. C. s. o. m. (2013) Value of sustainable reporting.

- Fama, E. F. and French, K. R. (1993) Common risk factors in the returns on stocks and bonds, *Journal of financial Economics* **33**, 3–56.
- Fama, E. F. and MacBeth, J. D. (1973) Risk, return and equilibrium: empirical tests, *Journal of Political Economy* **81**, 607–636.
- Finansdepartementet (2017) Ändrade regler för första–fjärde ap-fonderna (2000:192). lagen om allmänna pensionsfonder.
- Fjärde AP-fonden (n.d.) The ap-funds - a part of the pension system. URL: <http://www.ap4.se/en/about-ap4/the-ap-funds---a-part-of-the-pension-system/>. (accessed: 2018-02-22).
- 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**, 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**, 2646–2654.
- Generation (2012) Sustainable capitalism, *Generation Investment Management LLP*.
- Goyal, A. (2012) Empirical cross-sectional asset pricing: a survey, *Financial Markets and Portfolio Management* **26**, 3–38.
- Greene, W. H. (2003) *Econometric analysis*. New Jersey: Prentice Hall.
- Hong, H. and Kacperczyk, M. (2007) The price of sin: the effects of social norms in markets, *Journal of Financial Economics* **93**, 15–36.
- Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: implications for stock market efficiency, *The Journal of Finance* **48**, 65–91.
- Kempf, A. and Osthoff, P. (2007) The effect of socially responsible investing on portfolio performance, *European Financial Management* **13**, 908–922.
- Koerniadi, H., Krishnamurti, C., and Torani-Rad, A. (2013) Corporate governance and the variability of stock returns, *International Journal of Managerial Finance* **10**, 494–510.
- Limkriangkrai, M., Koh, S., and Durand, R. B. (2017) Environmental, social and governance (esg) profiles, stock returns and financial policy: australian evidence, *International Review of Finance* **17**, 461–471.

- Lintner, J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *The Review of Economics and Statistics* **47**, 13–37.
- MacKinlay, A. C. (1995) Multifactor models do not explain deviations from capm, *Journal of Financial Economics* **38**, 3–28.
- Manescu, C. (2011) Stock returns in relation to environmental, social and governance performance: mispricing or compensation for risk?, *Sustainable Development* **19**, 95–118.
- Markowitz, H. (1959) Portfolio selection: efficient diversification of investments. New York: Wiley.
- Mossin, J. (1966) Equilibrium in a capital asset market, *Econometrica* **34**, 768–783.
- MSCI (2018) Esg 101: what is esg investing. URL: <https://www.msci.com/esg-investing>. (accessed: 2018-05-15).
- Nasdaq (2017) Market cap segment review at nasdaq nordic exchanges. URL: <http://business.nasdaq.com/mediacenter/pressreleases/1660396/market-cap-segment-review-at-nasdaq-nordic-exchanges>. (accessed 2018-04-04).
- Petersen, M. A. (2009) Estimating standard errors in finance panel data sets: comparing approaches, *The Review of Financial Studies* **22**, 435–480.
- Przychodzen, J. and Przychodzen, W. (2014) Relations between eco-innovation and financial performance – evidence from publicly traded companies in poland and hungary, *Journal of Cleaner Production* **90**, 253–263.
- RobecoSAM (2017) Country sustainability ranking. URL: <http://www.robecosam.com/en/sustainability-insights/about-sustainability/country-sustainability-ranking/>. (accessed: 2018-04-04).
- Sharpe, W. F. (1964) Capital asset prices: a theory of market equilibrium under conditions of risk, *The Journal of Finance* **19**, 425–442.
- Sharpe, W. F. (1994) The sharpe ratio, *The Journal of Portfolio Management*. Available at: <http://web.stanford.edu/wfsharpe/art/sr/sr.htm>.
- Statman, M. (1987) How many stocks make a diversified portfolio?, *The Journal of Financial and Quantitative Analysis* **22**, 353–363.
- Statman, M. and Glushkov, D. (2009) The wages of social responsibility, *Financial Analysis Journal* **65**, 33–46.

Sveriges Riksbank (2018) Search interest & exchange rates. Data retrieved from Sveriges Riksbank. URL: <https://www.riksbank.se/en-gb/statistics/search-interest--exchange-rates/>.

Thomson Reuters (2018) Thomson reuters esg scores. URL: <https://financial.thomsonreuters.com/content/dam/openweb/documents/pdf/financial/esg-scores-methodology.pdf>.

Thomson Reuters Eikon (2018). (Database).

Appendix A. Grouping Technique

Beta values are calculated monthly based on five years of previous monthly data. If these estimates are used in a second regression, an errors in variables-problem occurs in the Beta values, which may result in an inconsistent estimate (attenuation bias) in the second regression. This is caused by the measurement error created in this first step, i.e. the estimated beta values contains both the true, unobservable beta values, plus measurement errors:

$$\hat{\beta}_{i,t} = \beta_{i,t} + v_{i,t} \tag{A1}$$

This error is assumed to be independent and identically distributed with zero mean. Thus, a so called grouping technique proposed by Fama and MacBeth (1973) can be used to reduce the variance of the errors, as the sample size approaches infinity, the variance approaches zero. The grouping technique can be summarized in two steps:

- Calculate the individual rolling betas of all the securities/stocks and order them in X portfolios based on these estimates.
- Calculate the portfolio equally weighted returns for each portfolio and regress them on the market excess returns, which gives the portfolio betas.

These betas have an error term that in the limit approaches zero as discussed previously, which removes the attenuation bias and gives consistent estimates in the Fama-MacBeth procedure. Fama and MacBeth (1973) creates 20 portfolios solely based on betas while Manescu (2011) does this with 50 portfolios. However, these are both ranked by beta and size. For a thorough derivation of this grouping methodology, see Manescu (2011) Fama and MacBeth (1973) and Black, Michael, and Scholes (1972).

Appendix B. Autoregressive models

Appendix B.1. ARCH

Time series have historically been studied in the context of homoscedastic processes, but evidence have been provided that for some kind of data the error variances $\text{var}[\varepsilon_t]$ appears in clusters of large and small forecast errors ε_t . This suggests heteroscedasticity where the variance of the forecast error depends on the previous error. This led to an alternative model called the autoregressive, conditional heteroscedasticity (ARCH) model. The ARCH model is as follows:

$$y_t = \beta' \mathbf{x}_t + \varepsilon_t$$

$$\varepsilon_t = u_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \quad (\text{B1})$$

where the first equation is a standard OLS regression in matrix notation and the second equality is ARCH estimation of it's error factor and u_t has a standard normal distribution. Furthermore, the expected value of the error conditioned on the vector \mathbf{x}_t and the previous value of the disturbance is assumed to be zero. In mathematical notation: $E[\varepsilon_t | \mathbf{x}_t, \varepsilon_{t-1}] = 0$, which makes the model a classical regression model. However, the error term is conditionally heteroscedastic not with respect to \mathbf{x}_t as is traditionally considered in regression models, but instead to ε_{t-1} . In mathematical notation:

$$\text{Var}[\varepsilon_t | \varepsilon_{t-1}] = E[\varepsilon_t^2 | \varepsilon_{t-1}] = E[u_t^2] [\alpha_0 + \alpha_1 \varepsilon_{t-1}^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Where ε_{t-1}^2 is the variance of the preceding disturbance.

If the process generating the disturbances inhibits certain properties (see Greene (2003, p. 241)), α_0 and α_1 is bound between zero and one, and it is the most efficient linear unbiased estimator of β . However, there exist a more efficient nonlinear estimator which is the maximization of the log-likelihood function which is used to estimate α_0 and α_1 in equation (B1) . The specification of the ARCH(1) becomes:

$$\sigma^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2. \quad (\text{B2})$$

In a more general format it can be written with q lags instead of 1:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2$$

In short, the ARCH(q) model is used to account for the heteroscedasticity of the disturbance which is caused by it's autocovariance.

Appendix B.2. GARCH

An extension to the ARCH(q) model is the generalized autoregressive conditional heteroscedasticity (GARCH(p,q)) which has the same underlying regression as in equation (B2), but it is conditioned on the information set at time t which is denoted by Ψ_t . The distribution of the disturbance is assumed to be normal: $\varepsilon_t | \Psi \sim N[0, \sigma_t^2]$, with the conditional variance:

$$\sigma_t^2 = \alpha_0 + \delta_1 \sigma_{t-1}^2 + \delta_2 \sigma_{t-2}^2 + \dots + \delta_p \sigma_{t-p}^2 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2$$

Where p and q denotes the number of lags in the respective terms. The intuition is that the conditional variance can be explained by two parts, the first part being it's autoregressive component of the error ε ,

where if ε_t is large ε_{t+1} should be large as well since α_1 is bound between 0 and 1. The second part is the autoregressive part of the variance σ^2 (Greene, 2003). One can say that the forecasted variance is a mix between the last periods forecast and the last periods disturbance. By only having a lag in each of the components the GARCH(p,q) model is simplified to a GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \delta_1 \sigma_{t-1}^2 + \alpha_1 \varepsilon_{t-1}^2 \quad (\text{B3})$$

Appendix C. The Charoenrook & Conrad methodology

Charoenrook and Conrad (2005) propose a method of identifying risk-based factors. They argue that there must exist linear relationship between the conditional mean and the conditional variance of the return on a factor-mimicking portfolio if the factor is a priced risk. They state that the expected excess return of asset i is a function of the risk-free rate and the covariance between the stochastic discount factor and the return of said asset:

$$E_t[R_{t+1}^i] - R^f = -R^f \text{cov}_t(M_{t+1}, R_{t+1}^i)$$

$E_t[R_{t+1}^i]$ is the return of asset i at time $t + 1$

R^f is the risk free rate

M_{t+1} is the stochastic discount factor at time $t + 1$

The intuition is that the excess return should be high if the covariance between the returns of the asset and the marginal utility of money is large and positive, if it is large and negative it should be the opposite. One can think of it as similar to the beta in CAPM, where a lower beta indicates lower risk and hence a lower risk premium. The problem is that the stochastic discount factor M is diffuse and hard to measure, but the authors extract M by assuming that it can be written as a linear function of risk factors.

$$M_{t+1} = \alpha_0 - \sum_{k=1}^K \alpha_k f_{t+1}^k + \eta_{t+1}$$

f^k has mean zero and η_{t+1} is white noise. A variable X may be related to some of these risk factors and can hence be explained by a function of them. Therefore, the return of asset i can be written as it's expected return plus its exposure to the risk factor X plus the exposure to all other risk factors f^k :

$$R_{t+1}^i = E_t[R_{t+1}^i] + \beta_t^{X^i} + \sum_{k=3}^K \beta_t^{ki} f_{t+1}^k \varepsilon_{t+1}^i$$

By sorting on X a factor mimicking portfolio is created by buying (selling) securities with high(low) values of X. By assuming that the factor mimicking portfolio is well diversified, both firm specific risk and variation due to the other risk factors f^k are diversified away. Evans and Stephen (1968) argue based on their findings that no more than ten securities are needed to create a well diversified portfolio, however Statman (1987) argues that the portfolio need to consist of at least 30 securities.

$$R_{t+1}^X - R^f = E_t[R_{t+1}^X - R^f] + B_t^X X_{t+1} \quad (C1)$$

Where B_t^X is a scaled parameter that the depends on the average beta with respect to the factor X, i.e. it's exposure to X. To analyze the excess return of the factor portfolio the relation between the stochastic discount factor M and X is examined by a projection of M onto X:

$$M_{t+1} = a_0 + \gamma X_{t+1} + v_{t+1}$$

Where v is orthogonal to X and related to the other risk factors f^k . By combining previous equations and simplifying them the excess return of the factor-mimicking portfolio is written as:

$$\begin{aligned} E_t[R_{t+1}^X] - R^f &= -R^f \text{cov}_t(\gamma X_{t+1} + v_{t+1}, R_t^X - R^f) \\ &= -R^f \text{cov}_t(\gamma X_{t+1} + v_{t+1}, E[R_{t+1} - R^f] + B_t^X X_{t+1}) \\ &= -R^f \gamma \text{cov}_t(X_{t+1}, E[R_{t+1}^X - R^f]) + B_t^X \gamma \text{cov}_t(X_{t+1}, X_{t+1}) \\ &\quad + \text{cov}_t(v_{t+1}, E[R_{t+1} - R^f]) + B_t^X \text{cov}_t(v_{t+1}, X_{t+1}) \end{aligned}$$

The first and third covariance term cancel out due to them having a constant (i.e. the expected value), while the forth term cancels out due to v_{t+1} being ortagonal to X_{t+1} . Since the covariance of the same term is the variance, the final equation becomes:

$$E_t[R_{t+1}^X] - R^f = -R^f \gamma B_t^X \text{Var}_t(X_{t+1}) \quad (C2)$$

This is the relation between the expected excess return and the variance of X itself, i.e. the excess return is explained by the variance of the risk factor. Furthermore, it can be shown that the conditional expected excess return of the factor mimicking portfolio should be linearly related to the portfolios conditional variance of the returns: By taking the variance of (C1):

$$\begin{aligned} \text{var}[R_{t+1}^X - R^f] &= \text{var}[B_t^X X_{t+1}] + \text{var}[E_t[R_{t+1}^X - R^f]] \\ &= (B_t^X)^2 \text{var}[X_{t+1}] \end{aligned}$$

This step is due to the basic properties of the variance. By dividing both sides with $(B_t^X)^2$:

$$\text{var}[X_{t+1}] = \frac{\text{var}[R_{t+1}^X - R^f]}{(B_t^X)^2}$$

Substituting this into (C2):

$$\begin{aligned} E_t[R_{t+1}^X] - R^f &= -R^f \gamma B_t^X \frac{\text{var}[R_{t+1}^X - R^f]}{(B_t^X)^2} \\ E_t[R_{t+1}^X] - R^f &= -R^f \frac{\gamma}{B_t^X} \text{Var}_t(R_{t+1}^X - R^f) \end{aligned} \quad (\text{C3})$$

"The intuition behind the relation is straightforward: the ex ante [as indicated by E_t] excess return, or compensation for risk, related to a factor should be related to the ex ante [as indicated by the Var_t] estimate of volatility for the factor in equilibrium, in proportion to how the factor covaries with the pricing kernel M " (Charoenruek and Conrad, 2005). Basically, if the variation related to the risk premium is constant, the risk-premium should also be constant, ceteris paribus. Furthermore, one should note that the parameter that relates the mean and the variance of the factor mimicking portfolio in (C3) must have the same sign as the conditional return of the same portfolio as indicated in (C2).

One interesting fact, given this model, is that the parameter that links the expected excess return and the conditional variance $-R^f \frac{\gamma}{B_t^X}$ is equal to Sharpe ratio divided by its standard deviation of return. Recall that the Sharpe ratio is defined as

$$\text{Sharpe ratio} = \frac{E[R^p - R^f]}{\sqrt{\text{var}[R^p - R^f]}}$$

This is simply done by dividing both sides of (C3) with the conditional variance and then multiplying by the standard deviation. The authors concludes that, ceteris paribus, when the expected volatility of the risk factor increases the expected excess return should scale up in the same fashion.

The method the authors use is first to test if there is a significant variation in the time series return of the factor mimicking portfolio, which is done by a GARCH(1,1) estimation. If the ARCH(1) and GARCH(1) term are significantly different from zero, it is concluded that there is significant time series variation. Then the relation between the conditional mean (portfolio excess return) and the conditional variance is tested by a simple linear regression:

$$R_{t+1}^X - R^f = \mu + \delta \sigma_{t+1}^2 + \eta_{t+1} \quad (\text{C4})$$

where σ_{t+1}^2 is calculated as proscribed in equation (ii). The interpretation is provided in section 4.2.

Appendix D. Thomson Reuters' ESG score

Environmental score	Descriptions
1. <i>Resource use</i> (20 measures, 11.2%)	Gives a company's performance and capacity to reduce the use of materials, energy and/or water. The company's ability to explore more eco-efficient solutions by improving the management of the supply chain.
2. <i>Emissions</i> (22 measures, 12.4%)	Gives the measures of the commitment and effectiveness of the company to reduce, environmental emissions in both the production and the operational processes.
3. <i>Innovation</i> (19 measures, 10.7%)	Gives the measures of the capacity of the company to reduce burdens and environmental costs for its customers. The reduction should create new market opportunities through development of new environmental technologies and processes or through eco-designed products.
Social score	
4. <i>Workforce</i> (29 measures, 16.3%)	Gives the company's effort to sustain job satisfaction, a healthy and safe workplace, maintaining diversity and equal opportunities, as ability to create development opportunities for its employees.
5. <i>Human rights</i> (8 measures, 4.5%)	Gives the efficiency of the company to respect the fundamental human rights conventions.
6. <i>Community</i> (14 measures, 7.9%)	Gives the commitment of the company being good citizens, protecting public health and, respecting business ethics.
7. <i>Product responsibility</i> (12 measures, 6.7%)	Gives the company's capacity to produce goods and services of quality and the ability to integrate the customer's health and safety, integrity and data privacy.
Governance score	
8. <i>Management</i> (34 measures, 19.1%)	Gives the,commitment and efficiency towards using the best practice corporate, principles.
9. <i>Shareholders</i> (12 measures, 6.7%)	Gives the commitment and efficiency of a company towards its shareholders and the, practise of anti-takeover tools.
10. <i>CSR strategy</i> (8 measures, 4.5%)	Gives the way the company communicates the economic, social and environmental, dimensions into the daily decisions-making processes.

% gives the percentage of the aggregated ESG score which contains totally 178 individual scores (Thomson Reuters (2018)).

Appendix E. List of ESG Companies

D Carnegie & Co AB	Fabege AB	Saab AB
Oriflame Holding AG	Hexagon AB	SSAB AB
Capio AB (publ)	Investor AB	Swedish Match AB
Atlas Copco AB	Nordea Bank AB	Trelleborg AB
Swedish Orphan Biovitrum AB (publ)	NCC AB	Telefonaktiebolaget LM Ericsson
Castellum AB	Getinge AB	Nolato AB
Beijer Ref AB (publ)	Intrum Justitia AB	Hexpol AB
Axfood AB	Gunnebo AB	Boliden AB
Fastighets AB Balder	Husqvarna AB	Kindred Group PLC
Bergman & Beving AB	H & M Hennes & Mauritz AB	Alfa Laval AB
Assa Abloy AB	Nederman Holding AB	Skandinaviska Enskilda Banken AB
Elekta AB (publ)	Nibe Industrier AB	AB SKF
Clas Ohlson AB	Holmen AB	Telia Company AB
CTT Systems AB	Kinnevik AB	Lundin Petroleum AB
ICA Gruppen AB	SAS AB	Electrolux AB
Fingerprint Cards AB	Svenska Handelsbanken AB	Svenska Cellulosa SCA AB
JM AB	Swedbank AB	Sandvik AB
Lindab International AB	Skanska AB	Volvo AB
Kungsleden AB	Svedbergs i Dalstorp AB	Tele2 AB
Nobia AB	Ratos AB	
BillerudKorsnas AB (publ)	SECTRA AB	
L E Lundbergforetagen AB (publ)	Securitas AB	
Modern Times Group MTG AB	Hufvudstaden AB	
Eniro AB	Industrivarden AB	
	VBG Group AB (publ)	
	Wihlborgs Fastigheter AB	

Above is the companies from the Swedish market with a Thomson Reuter ESG score for Jan 2003-Dec 2017.

Appendix F. Fama-MacBeth in Stata

As the Fama Macbeth methodology is not a part of the Stata code, it can be rather tricky to conduct. There are however good packages provided by the Boston College Statistical Software Components (SSC). It is important to realize that these packages require the beta values to be calculated beforehand.

Appendix F.1. Beta calculation

Before initializing the Fama Macbeth procedure, the beta values need to be calculated by a regression on the asset returns on the market excess returns in a traditional CAPM style. The total number of regressions becomes $N \times T$, where N is the number of stocks and T is the number of time units. The approach is as follows:

- Import the necessary data in long format.
- `xtset` the data with the id and time variables.
- Run the regressions, the `rangestat` function is recommended.
- Replace estimates by missing values if there is insufficient data points.
- Save and append to the main dataset

However, as described in the section appendix A, there is an error in variables problem which should be taken care of by a grouping technique. This is taken care of as follows:

- Use the betas estimated from the previous step to generate X portfolios.
- Use the generated portfolios to create portfolio mean returns.
- Regress the portfolio returns on the market excess return.
- Assign the portfolio betas to the corresponding stock

Appendix F.2. The Fama Macbeth regression

The Fama Macbeth regression can be made in four different ways:

1. By own code, where the `statsby` function is recommended for the regression and `foreach` loops are recommended for the calculation of t-statistics
2. By the `Asreg` package⁶
3. By the `xtfmb` package⁴
4. By the `fm.do` files⁷

⁶Available by the `findit` function in Stata

⁷Available at http://www.kellogg.northwestern.edu/faculty/petersen/htm/papers/se/se_programming.htm

Appendix G. Additional tables and graphs

Without industry dummies									
Variables	Jan 2003-Dec 2017			Jan 2003-Dec 2008			Jan 2009-Dec 2017		
Beta	0.114 (0.712)	0.164 (0.604)	0.316 (0.433)	-0.172 (0.794)	-0.050 (0.942)	0.197 (0.819)	0.235 (0.515)	0.249 (0.495)	0.383 (0.455)
Size	-0.192 (0.178)	-0.239 (0.119)	-0.319* (0.074)	-0.280 (0.299)	-0.355 (0.244)	-0.587* (0.085)	-0.143 (0.341)	-0.169 (0.256)	-0.130 (0.438)
Book-to-market	0.006 (0.976)	0.004 (0.984)	-0.054 (0.808)	-0.101 (0.448)	-0.115 (0.373)	-0.128 (0.591)	0.182 (0.609)	0.178 (0.612)	0.082 (0.826)
Momentum	0.121 (0.108)	0.114 (0.136)	0.171** (0.017)	0.022 (0.774)	0.015 (0.867)	0.107 (0.356)	0.174 (0.122)	0.173 (0.120)	0.204** (0.026)
ESG (1-10)	0.001 (0.847)			0.004 (0.798)			0.004 (0.617)		
Environmental (1-3)		-0.002 (0.732)			0.004 (0.713)			-0.006 (0.397)	
Social (4-7)		0.011* (0.051)			0.011 (0.295)			0.012* (0.062)	
Governance (8-10)		-0.007* (0.066)			-0.012* (0.084)			-0.001 (0.712)	
1. Resource use			0.003 (0.466)			0.004 (0.604)			0.003 (0.662)
2. Emissions			-0.004 (0.474)			-0.004 (0.716)			-0.006 (0.186)
3. Innovation			0.000 (0.943)			0.008 (0.281)			-0.002 (0.843)
4. Workforce			0.003 (0.644)			-0.001 (0.963)			0.009 (0.206)
5. Human rights			-0.006 (0.320)			0.011 (0.129)			-0.018*** (0.007)
6. Community			0.006* (0.092)			0.007 (0.278)			0.008* (0.072)
7. Product responsibility			0.004 (0.333)			0.009 (0.216)			-0.002 (0.610)
8. Management			-0.010** (0.022)			-0.017** (0.047)			-0.004 (0.179)
9. Shareholders			0.005* (0.096)			0.011* (0.066)			0.003 (0.394)
10. CSR strategy			0.004 (0.256)			-0.001 (0.792)			0.007 (0.166)
Constant	5.412 (0.101)	6.503* (0.069)	8.389** (0.032)	7.725 (0.236)	9.433 (0.202)	13.618* (0.087)	3.857 (0.243)	4.488 (0.168)	4.336 (0.204)
Observations	7,809	7,809	7,809	2,378	2,378	2,378	5,430	5,430	5,430
Groups	180	180	180	72	72	72	108	108	108
Average R ²	0.246	0.291	0.449	0.261	0.318	0.524	0.237	0.274	0.398

Table X: Equation (iv) estimated with the Fama-Macbeth Procedure for the whole time period and the sub periods, without controlling for industry sectors. The returns (dependent variable) are in percentage form. P-values in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Beta values estimated without the grouping technique. Groups is the number of time intervals (months).

With industry dummies									
Variables	Jan 2003-Dec 2017			Jan 2003-Dec 2008			Jan 2009-Dec 2017		
Beta	-0.162 (0.656)	-0.040 (0.916)	-1.259 (0.449)	-0.923 (0.126)	-0.936* (0.076)	-3.275 (0.427)	0.130 (0.714)	0.382 (0.357)	0.472 (0.335)
Size	0.011 (0.951)	0.006 (0.972)	0.462 (0.304)	0.039 (0.897)	0.038 (0.882)	0.930 (0.365)	-0.083 (0.727)	-0.098 (0.705)	0.061 (0.824)
Book-to-market	-0.078 (0.806)	-0.102 (0.750)	-0.319 (0.501)	-0.340 (0.379)	-0.403 (0.350)	-0.051 (0.954)	0.037 (0.933)	0.042 (0.921)	-0.034 (0.936)
Momentum	0.156* (0.097)	0.155* (0.080)	-0.215 (0.517)	0.066 (0.506)	0.051 (0.520)	-0.695 (0.362)	0.200 (0.167)	0.216 (0.118)	0.240** (0.044)
ESG (1-10)	-0.007 (0.431)			-0.008 (0.543)			0.001 (0.953)		
Environmental (1-3)		-0.004 (0.637)			0.009 (0.481)			-0.014 (0.199)	
Social (4-7)		0.002 (0.673)			-0.009 (0.238)		0.016** (0.035)		
Governance (8-10)		-0.004 (0.432)			-0.008 (0.419)		0.000 (0.990)		
1. Resource use			-0.020 (0.469)			-0.063 (0.289)			0.002 (0.844)
2. Emissions			-0.014 (0.245)			-0.025 (0.407)			-0.005 (0.285)
3. Innovation			0.006 (0.606)			0.039** (0.028)			-0.010 (0.326)
4. Workforce			0.017 (0.332)			0.029 (0.506)			0.006 (0.314)
5. Human rights			-0.021* (0.051)			-0.012 (0.576)			-0.018*** (0.001)
6. Community			-0.001 (0.869)			-0.013 (0.458)			0.010** (0.018)
7. Product responsibility			0.020 (0.278)			0.057 (0.172)			-0.003 (0.148)
8. Management			-0.019** (0.034)			-0.028 (0.115)			-0.006*** (0.008)
9. Shareholders			-0.007 (0.664)			-0.019 (0.622)			0.005 (0.258)
10. CSR strategy			-0.002 (0.725)			-0.024 (0.274)			0.005 (0.324)
Constant	1.176 (0.766)	1.592 (0.674)	0.733 (0.885)	0.911 (0.901)	2.286 (0.692)	3.626 (0.706)	2.672 (0.590)	2.672 (0.635)	-0.819 (0.891)
Observations	7,809	7,809	7,809	2,378	2,378	2,378	5,430	5,430	5,430
Groups	180	180	180	72	72	72	108	108	108
Average R ²	0.536	0.571	0.694	0.630	0.674	0.816	0.472	0.500	0.611

Table XI: Equation (iv) estimated with the Fama-Macbeth Procedure for the whole time period and the sub periods, controlling for industry sectors. The returns (dependent variable) are in percentage form. P-values in parenthesis, *** p<0.01, ** p<0.05, * p<0.1. Beta values estimated with the grouping technique with 10 portfolios. Groups is the number of time intervals (months).

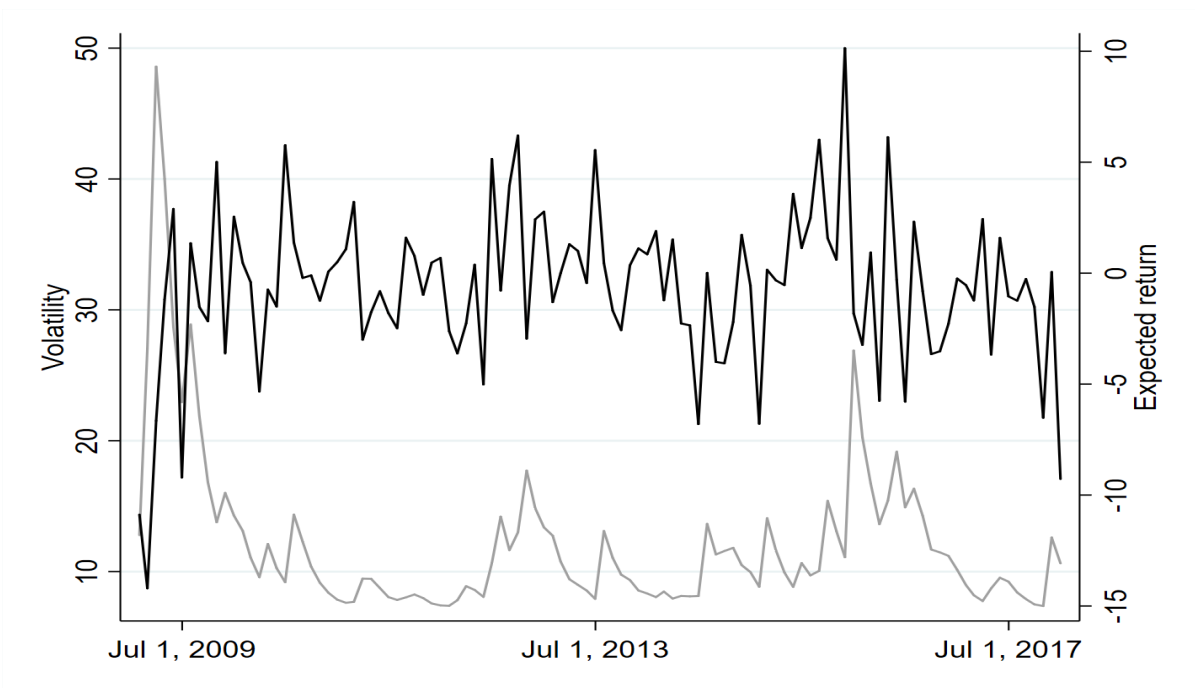


Figure 1: Conditional mean (Expected return, $R_{t+1}^{social} - R^f$, indicated by the black line) and conditional variance (Volatility, σ^2 , indicated by the grey line) of the *social* risk factor mimicking portfolio constructed as indicated in equation (vii) for the time period of Jan 2009-Dec 2017

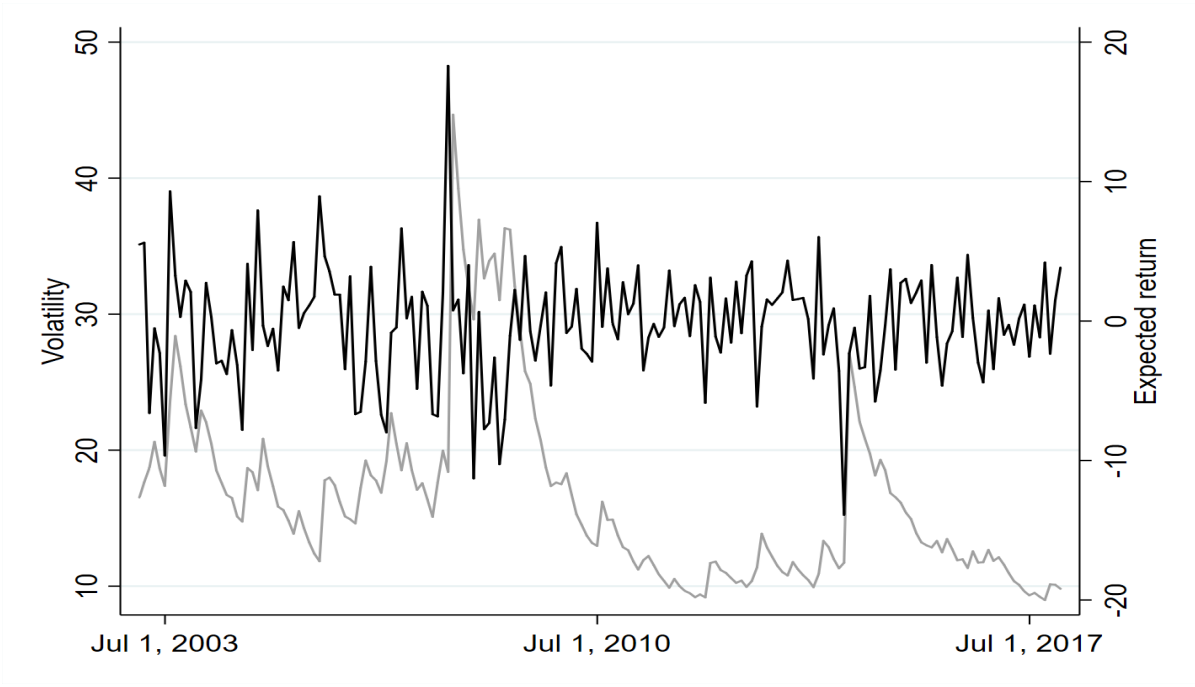


Figure 2: Conditional mean (Expected return, $R_{t+1}^{humanrights} - R^f$, indicated by the black line) and conditional variance (Volatility, σ^2 , indicated by the grey line) of the *human rights* risk factor mimicking portfolio constructed as indicated in equation (vii) for the time period of Jan 2003-Dec 2017

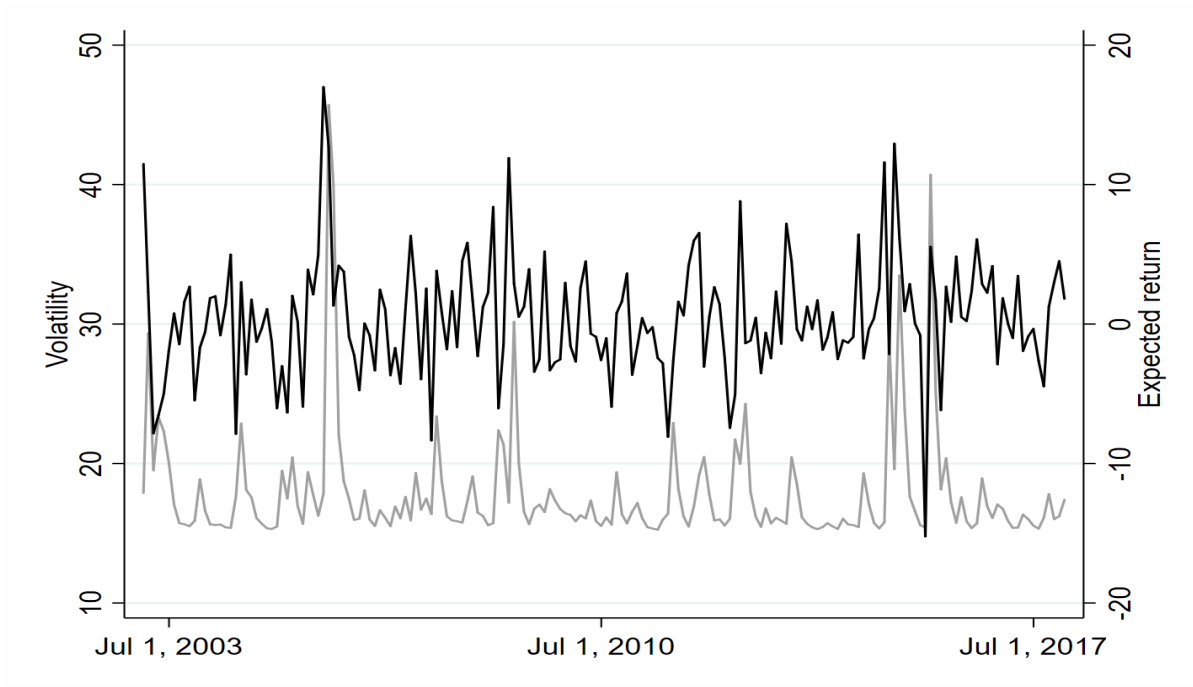


Figure 3: Conditional mean (Expected return, $R_{t+1}^{management} - R^f$, indicated by the black line) and conditional variance (Volatility, σ^2 , indicated by the grey line) of the *management* risk factor mimicking portfolio constructed as indicated in equation (vii) for the time period of Jan 2003-Dec 2017.