Universitetet
i Stavanger

UIS BUSINESS SCHOOL

MASTER'S THESIS

STUDY PROGRAM: Master of Science in Business Administration	THESIS IS WRITTEN IN THE FOLLOWING SPECIALIZATION/SUBJECT: Applied Finance IS THE ASSIGNMENT CONFIDENTIAL? **(NB!** Use the red form for confidential theses)

TITLE:
Focus on sustainability in the oil and gas industry – increasing investors' risk-adjusted return, or harmful to their wealth?

AUTHOR(S)		SUPERVISOR:
Candidate number:	Name:	Christian Jensen
3116 …………………	Are Rettedal Ekeli ………………………………………	
…………………	………………………………………	

Abstract

Expectations of increased global energy demand in the years to come, together with a desperate need to decrease global emissions, have contributed to increased investor attention to the environment and sustainability. This research explores the implications of this focus on investors' risk-adjusted return.

The strong link between negative environmental externalities and the products the oil & gas industry produce and sell make this industry particularly relevant. Clear differences among the firms in the industry concerning their role and responsibility towards the reduction of emission make the industry even more interesting. Therefore, this research seeks to answer the following research question:

Focus on sustainability in the oil and gas industry – increasing investors' risk-adjusted return, or harmful to their wealth?

Sustainability performance is measured through ESG-scores, provided by MSCI. Using panel data models, this research fails to find any significant relationship between ESG performance and stock return. In addition, the effect of changed ESG-score, called momentum, was tested on both stock return and stock volatility, but the results were inconclusive.

However, this research finds that firms with good ESG performance have significantly lower monthly stock volatility than firms with poor ESG performance. The reduction is slightly above one-fifth of the poor performers' volatility. Several possible explanations are presented, including a lower likelihood for firms with good ESG performance to experience scandals and receive negative press publicity. Another explanation is that the good ESG performers are more likely to invest in renewable energy, which reduces their exposure to oil & gas prices. Hence, more diversified sources of revenues could explain the lower stock volatility for these firms.

Investors who seek wealth maximization through the highest possible risk-adjusted return is, based on these results, recommended to invest in oil and gas companies with good ESG performance.

Foreword

The submission of this master thesis completes my Master of Science in Business Administration with a specialization in Applied Finance, at the University of Stavanger.

A strong personal interest in the role investors and finance can play in the attempt to solve the paradox of increasing demand for energy while reducing toxic emissions laid the foundation for this research within the oil & gas industry.

The process of writing this thesis has been challenging but also given me tools and knowledge that I can utilize in my future career.

I would like to thank my supervisor, Christian Jensen, for his feedback and guidance throughout the whole semester. I am grateful for your availability and quick replies.

Finally, I would like to thank my partner, Karoline, and my daughter Olivia, for their patience and support.

Stavanger, June 14th, 2020

Are Rettedal Ekeli

Table of contents

Abstract ii

Foreword iii

List of tables v

List of figures vi

List of abbreviations vi

1. Introduction 1
- **1.1 Background and context** 1
- **1.2 Research Question** 2
- **1.3 Scope and structure of the research** 3

2. Previous Research 4
- **2.1 Concept and research on ESG** 4
- **2.2 Ways of incorporating sustainability into investments decisions** 7
- **2.3 Shareholder and stakeholder position** 8

3. Method 9
- **3.1 Construction of the dataset** 9
  - 3.1.1. Subindustries in the research 9
  - 3.1.2 Choice of firms 10
  - 3.1.3 Choice of sustainability index and agency 10
  - 3.1.4 Factors controlled for in the analysis 12
  - 3.1.5 Other collection of data 15
- **3.2 Panel data method and models** 15
  - 3.2.1 Advantages of panel data 16
  - 3.2.2 The different panel data models 16
  - 3.2.3 Key assumptions 18
  - 3.2.4 Tests of suitability 18

4. Data and statistics 20
- **4.1 The effect of ESG performance on stock returns - statistics** 20
- **4.2 The effect of ESG-score on stock volatility – statistics** 24

5. Analysis 25
- **5.1 The effect of ESG performance on stock returns - analysis** 26
  - 5.1.1 F-test 26
  - 5.1.2 Breusch-Pagan Lagrange Multiplier-test 26
  - 5.1.3 Hausman-test 27
  - 5.1.4 Fixed effects-model 27
  - 5.1.5 Pooled model 30
- **5.2 The effect of ESG performance on stock volatility- analysis** 32
  - 5.2.1 F-test, BP LM-test and Hausman-test 33
  - 5.2.2 Fixed effects model 34
  - 5.2.3 Random-effects model 37

6. Discussion of results 40
- **6.1 The effect of ESG performance on stock returns** 40
  - 6.1.1 Possible explanations 41
  - 6.1.2 Possible implications 42
  - 6.1.3 The effect of ESG momentum on stock returns – possible explanations 43
- **6.2 The effect of ESG performance on stock volatility** 43
  - 6.2.1. Possible explanations 44

6.2.2 Possible implications ... 45
6.2.3 The effect of ESG momentum on stock volatility – possible explanations ... 46

7. A critical view, limitations, and further research ... 47

8. Conclusion ... 49

Bibliography ... 50

Appendices ... 53
Appendix 1 ... 53
Appendix 2 ... 54
Appendix 3 ... 55
Appendix 4 ... 56
Appendix 5 ... 57
Appendix 6 ... 58

List of tables

Table 1 - summary statistics and characteristics of variables in dataset (1) with stock return as the dependent variable. ... 21
Table 2 - summary statistics and characteristics of variables in dataset (2) with stock volatility as the dependent variable. ... 24
Table 3 - Summary of results of model 1A in dataset (1) with stock return as dependent variable. ... 27
Table 4 - Summary of results of model 1A in dataset (1). ... 28
Table 5 - Summary of results of model 1B in dataset (1) with momentum as explanatory variable. ... 28
Table 6 - Summary of results of model 2A in dataset (1) with stock return as dependent variable. ... 31
Table 7 - Summary of results of model 2A in dataset (1). ... 31
Table 8 - Summary of results of model 2B in dataset (1) with momentum as explanatory variable. ... 31
Table 9 - Summary of results of model 1A in dataset (2) with stock volatility as the dependent variable. ... 34
Table 10 - Summary of results of model 1A in dataset (2). ... 35
Table 11 - Summary of results of 1B in dataset (2) with momentum as explanatory variable. ... 35
Table 12 - Summary of results of model 2A in dataset (2) with stock volatility as dependent variable. ... 38
Table 13 - Summary of results of model 2A in dataset (2). ... 38
Table 14 - Summary of results of model 2B in dataset (2) with momentum as explanatory variable. ... 39

List of figures

Figure 1 - Factors considered in MSCIs' assessment of ESG-score. 5
Figure 2 - Fraction of companies in each subindustry. 9
Figure 3 - Frequency of different homecountries for the firms in the research. 10
Figure 4 – The sample distribution of ratings, both original ratings from MSCI, and the broader categories used in this research.. 12
Figure 5 - industry return and oil price change. 13
Figure 6 - Industry volatility and oil price volatility. 13
Figure 7 - Number of upgraded and downgraded ESG-scores each year. 23
Figure 8 - Distribution of estimated individual-specific effects. 37

List of abbreviations

ESG: Environmental, social and governance

SRI: Social responsible investing

CSR: Corporate social responsibility

E&P: The subindustry exploration and production. Also known as the upstream segment

RMTS: The subindustry refining, marketing, transportation and storage. Also known as the downstream segment

Integrated: The subindustry integrated oil & gas

EBITDA-margin: Fraction of earnings before interest, tax, depreciation, and amortization of total revenues

RE model: Random effects model

FE model: Fixed effects model

BP LM-test: Breusch-Pagan Lagrange Multiplier- test

ESG momentum: The event of upgraded or downgraded ESG-score

1. Introduction

1.1 Background and context

Back in the late 1970s and early 1980s, Exxon Mobile's research, both external and internal, concluded that the production of non-renewable energy resources such as oil and gas would lead to dramatic climate change effects if hard actions were not taken quickly. Exxon responded by ignoring this threat and instead stood in front of climate change denial (Banerjee, Song, & Hasemyer, 2015).

Ever since more and more research has been conducted every year and the message is clear; world energy consumption is projected to rise by 50% between 2018 and 2050 (U.S. Energy Information Administration, 2019, p.23), but emissions need to decrease drastically to prevent catastrophic situations in the future. This includes dramatic effort from the oil and gas industry which produces the fuel types that is responsible for about 55% (in 2017) of total CO_2 emissions in the world (Ritchie & Roser, 2017).

Attention from stakeholders has grown with the increased threat of dramatic climate changes. This, of course, includes environmental organizations, but also employees, government, and, as this research focuses on; Investors. Already back in 2001, one of the firms involved in this research, Equinor (then Statoil), released its first sustainability yearbook (Statoil, 2001, p.4), as an answer to the increased interest in the environment, climate and corporate social responsibility from investors (Borchgrevink, 2019, p.289).

Some of the firms in the industry have responded to the challenge by stating that they will continue as before, do what they do best (exploit and use their core competencies), and let investors diversify any potential extra risk themselves. This is the case for Aker BP, which will continue to invest in oil and gas projects. If investors want renewable energy sources in their portfolio, they can invest in such assets themselves (Degnes, 2020).

On the other hand, firms like Equinor, Royal Dutch Shell, and Total wants to react to the changing times and increase their efforts to reduce emissions and invest in renewable energy resources (Degnes, 2020).

With such conflicting attitudes and beliefs to the challenge of sustainability across the industry, it is interesting to test which strategy has been the most successful from an investor's point of view. Do the deviations from core competencies, and other related costs of sustainability focus lead to weaker performance, and as a consequence, reduced stock returns? Or will the focus on sustainability be a source of competitive advantage, increased popularity, and higher stock returns?

Another important aspect from an investor's perspective is the volatility of the stock prices, a general measure of the total risk of the different stocks. Will the risk differ based on the focus and performance the firms have related to sustainability?

Literature and research related to these questions finds mixed or inconclusive results, and are generally concerned with the whole market and not specific industries. This research seeks to expand the scope of the literature and provide updated results by focusing specifically on an industry that is highly responsible for the world's emissions, and by using recent data.

This research use ESG-scores, which are the quantitative measures of the concept ESG, to test the effect of sustainability. It is a tool and a concept that allows investors to incorporate sustainability and corporate responsibility into their investment decisions (Nilsen & Noergaard, 2011). ESG-scores are developed by rating agencies and are determined through the evaluation and ranking of the firm's performance within environmental, social, and governance (which ESG is the acronym of) related factors (MSCI, 2020).

The ESG-scores will be the primary tool in the attempt to find out if the focus on sustainably in the industry is a source to superior risk-adjusted return for investors, or if it is harmful to their wealth.

1.2 Research Question

The purpose of this research project is to test whether the focus on sustainability by firms in the oil and gas industry, gives different (better or worse) value to investors in terms of stock returns and volatility.

Hence, the following research question is formulated;

The question will be examined through a series of statistical tests using panel data models. To be able to conclude on the research question, the effect of sustainability focus on both stock return and stock volatility is tested.

1.3 Scope and structure of the research

The research's scope

This research is limited to the oil and gas industry. ESG data is collected from MSCI, and only firms listed by this rating agency are included in the study. The data are from the start of May 2015 to the end of December 2019. The starting period is May 2015 because it is the start of the publicly available data on ESG ratings from MSCI (for the oil and gas industry). The end of December 2019 was chosen as the deadline for data collection because the collection of data started in early January 2020. The choice of data and variables are further explained in chapter 3, including data providers for the different variables. The collection of the dataset is the independent work of this author.

Structure of the research

The research is divided into eight chapters of varying scope and purpose. The chapters include subsequent subchapters.

Chapter 1 provides an introduction to the topic and puts the research into context. In this chapter, the research question is also presented.

Chapter 2 looks backward and considers existing research and the concept of sustainable investing. In chapter 3, the chosen research design is presented. The chapter covers the construction of the dataset, and the method of panel data, including necessary assumptions and tests of suitability.

Chapter 4 presents the basic statistics and features of the different variables before the actual analysis, and the results are covered in chapter 5.

Chapter 6 provides a discussion of the results from the analysis, including possible explanations and the implications for investors.

Chapter 7 presents a critical view of the research and discuss how further research could build on the findings of this research. Chapter 8 summarize the research and conclude on the research question.

2. Previous Research

This chapter look backward and provides the foundation for the analysis and research in the sections to come. The attention is brought to sustainable investing and, in particular, ESG. This includes both explanations of the concepts as well as findings of previously conducted research.

2.1 Concept and research on ESG

CSR (corporate social responsibility), SRI (social responsible investing), and ESG (environmental, social, and governance) are similar concepts, with the same underlying purpose of highlighting the responsibility firms and investors have towards sustainability. The concepts are often mixed, but the main difference is that CSR is related to the firm's management, and its long term strategic plans to create value and competitive advantages. At the same time, ESG is a concept related to investors, and their desire to incorporate sustainability factors into their valuation models (Nilsen & Noergaard, 2011). SRI is also focusing on investors but differs from ESG because SRI screens investment opportunities, not only to incorporate these factors but to eliminate (invest in) poor (good) sustainability performing firms (Zhou, 2019).

This research is concerned with the implications of investor focus (or lack of focus) on sustainability in the oil & gas industry and the different performance of the companies. Therefore, ESG is the preferred concept to use as evaluation criteria to explore these differences.

A lot of factors and data are evaluated to end up with a general ESG-score. Figure 1 below shows which factors MSCI (which is the chosen ESG-ratings provider of this paper) considers when they measure the ESG performance of the different firms. They use publicly available data to collect over 1000 different data points to conduct the assessments on the below factors (MSCI, 2020).

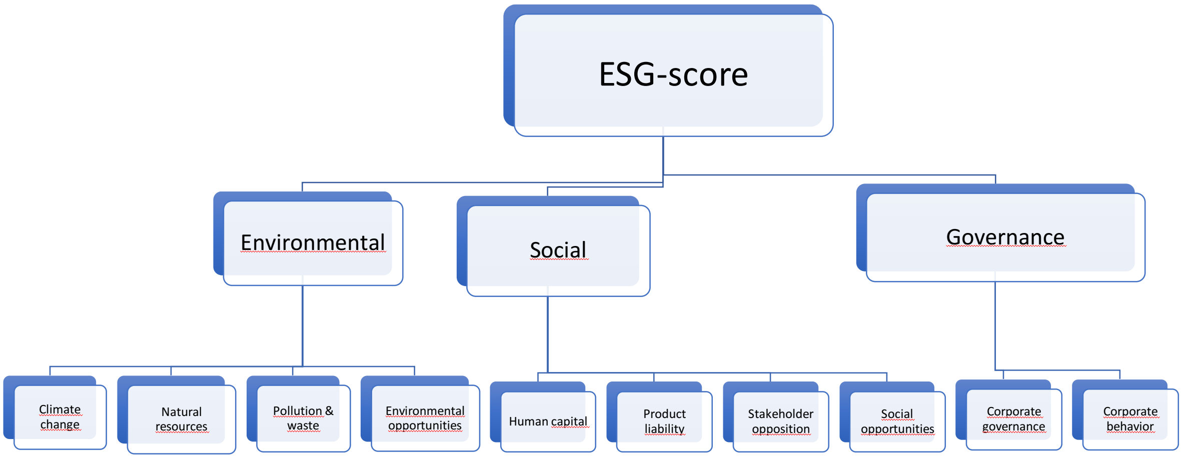

Figure 1 - Factors considered in MSCIs' assessment of ESG-score. Source: Based on MSCI (2020) presentation of considered factors.

The categories at the bottom of figure 1 are considered in further detail to measure the performance in each category. For example, to measure the performance concerning the category “human capital” (below “social”), MSCI will go one level further down and consider key ESG factors like labor management, health and safety, human capital development, and supply chain labor standards. In total, 37 different key ESG factors are examined and evaluated (MSCI, 2020).

After evaluating the risk associated with the key ESG factors, these are assigned with weights based on the perceived impact they impose. Finally, these weights and risks are combined and standardized relative to industry peers to develop the final ESG-score for a given company (MSCI, 2020).

A substantial amount of research investigates the effect of ESG-ratings on financial performance. The research has not led to any consensus yet as it differs on both methodologies, objectives, and conclusions. MSCI has conducted internal research, through Giese and Lee (2019), where they aim at consolidating the findings on the topic. They suggest that one of the reasons for the mixed results is the relatively short existence of ESG ratings, which implies that more extended time series are needed to provide consistency in the results.

They find that companies with high (good) MSCI ESG rating, historically have a lower frequency of major ESG related risk incidents, and are more adaptable to changing market conditions and environments (which they refer to as lower systematic risk). Another finding is that high ESG rated firms were more profitable than their peers. However, they find mixed evidence of the rating's effect on stock performance, including differences among regions.

Also, on behalf of MSCI, Giese and Nagy (2018) investigated the effect of momentum (changed ESG rating) in developed markets, and found that upgraded (downgraded) ESG scores led to higher (lower) stock returns.

Using MSCI ESG KLD STATS, the precursor to the current MSCI ratings, Bansal, Wu, and Yaron (2016) manages to expand the time series using data from 1991 to 2011. They found that companies with good performance within ESG factors delivered significantly higher alphas (risk-adjusted return compared to a benchmark) than the companies that have poor performance regarding ESG. The effect is, however, temporary and varying over time. They also find evidence of a more considerable alpha difference between the firms in cyclical upturns than during recessions, which, they claim, shows that investors' preferences for stocks with good performance within ESG factors behave similarly to luxury goods.

In contrast, Marsat and Williams (2011) find evidence of the opposite, namely a negative effect of responsible behavior on stock returns in the period 2005 to 2009. However, the methodology differs substantially in terms of the period, data point frequency, and control variables. Another issue is the way different studies code the ratings before putting them into the econometrics software. Such methodology issues represent a general problem in the research and a likely explanation to the differences in results.

Results from ESG research also quickly becomes less relevant. Studies from as recently as 2010 are not necessarily representative of the situation in 2020, as both investor preferences and awareness regarding ESG, changes rapidly. The rating process and data availability are also continuously improving.

Besides, most studies are concerned with the effect of ESG on the market as a whole and not specific industries. In comparison, this research specifically tests the effect of ESG on the oil and gas industry. This is particularly interesting because of the strong link between negative environmental externalities and the commodity the oil and gas industry produce and sell. The clear strategic crossroad (continue as before and exploit core competencies versus increased efforts to reduce emissions and invest in renewable energy) within the industry, makes the industry even more relevant to research.

2.2 Ways of incorporating sustainability into investments decisions

There are multiple ways ESG performance could affect investors' decisions. One way to look at it is through a premium in the cost of capital estimations for firms with poor ESG performance. PWC conducts an annual analysis of the risk premium in the Norwegian markets, where they receive information from members of the Norwegian society of financial analysts. Results from this analysis in 2019 show that 57% of the respondents agree that a risk premium for poor performance related to environment and sustainability should be included in the cost of capital when evaluating companies, up from 54% in 2018 (PwC, 2019, p.16).

Some respondents also commented that environment and sustainability could be adjusted for in the cash flow estimates instead of the cost of capital (PwC, 2019, p.16). In both cases, the main objective is to incorporate these factors into the estimated value of the investment.

Weber (2008) argues for a procedure where both the benefits and costs related to CSR are incorporated into the investment decision. The method is called monetary CSR Value Added and uses the well-known discounted cash flow approach to find the present value of the monetary CSR Value Added.

$$Monetary\ CSR\ Value\ Added = \sum_{n=1}^{n=\infty} (B_n^{CSR} - C_n^{CSR}) * \frac{1}{(1+i)^n}$$

Where n represents periods, and i is the discount rate. B_n^{CSR} describes the benefits of CSR, while C_n^{CSR} Incorporates the costs of CSR (Weber, 2008).

Though it seems complicated to identify these, often qualitative, benefits, and costs, Weber argues that the firms should seek to identify quantitative indicators. She provides examples such as repurchase rates, market share, hiring rates and reputation indices and rankings.

CSR is related to the firm’s management. Still, this procedure could also relate to investors seeking to incorporate sustainability into their assessment and valuation of different firms and industries.

2.3 Shareholder and stakeholder position

The concepts or positions of shareholder value and stakeholder society is probably the broadest discussion and include perhaps the most distinctive differences in perspective on how firms and their management should act concerning sustainability and social responsibility.

The thought that individual egoistic behavior can lead to a favorable outcome for society as a whole can be traced back to the classical economist Adam Smith, which was a strong proponent of market liberalism (Dedekam Jr, 2002, p.60).

Such a view is consistent with what Tirole (2006, p.56-62) refers to as the shareholder value position, which claims that the ultimate aim for a firm and its management is to maximize shareholder wealth and that the prices will reflect the scarcity of resources.

Proponents of this view claim that the best way to deal with externalities that arise from the activities the firm conducts is through legal and contractual actions. Creditors and employees should secure their interest through contracts and legal protection, for example, to prevent expropriation of creditors and lack of safety on the job for employees (Tirole, 2006, p.56-62)

In contrast, the stakeholder position, which Tirole (2006, p.56-62) refers to it as, claims that the target of corporations is to act responsibly and to fulfill a larger purpose to society. This is done by considering other stakeholders like employees, communities, creditors, and in general, take ethical considerations into account when making decisions. Management should hence include externalities into their analysis when they consider different possible choices. They should also consider the impact their decisions have on the environment, even if this leads to reduced profits.

With the assumption that it is commercially profitable to take different stakeholder's interests into account, it would be possible to maximize shareholder value by taking ethical considerations and responsibility. Tirole (2006, p.56-62) refers to situations where firms implement employee rights beyond the standard may help them attract top talented individuals. The same reasoning can apply to suppliers. A similar logic can be utilized on ethical considerations, which are likely to increase short-term costs, but these costs may be outweighed by value increasing long term benefits. Treating stakeholders fair could hence be a source of competitive advantage and superior financial performance.

However, the underlying intentions of the management may, rightfully, be questioned. To utilize the stakeholder society to gain shareholder value would be classified as the shareholder value position.

3. Method

This chapter is divided into two parts, where the first part considers the necessary choices and methods to construct the dataset for the research. The second part will present the method of panel data analysis (statistical tests) in which the dataset is processed through to find answers to the research question.

3.1 Construction of the dataset

This part goes through the variables included in the dataset. Because the dataset was constructed for this research specifically, this part provides information on the included variables, why they were included, and how they were collected. This will make it possible for others to replicate the dataset.

3.1.1. Subindustries in the research

The firms in this research belong to different subindustries within the oil and gas industry. The shares in figure 2 show the fraction of companies in each subindustry.

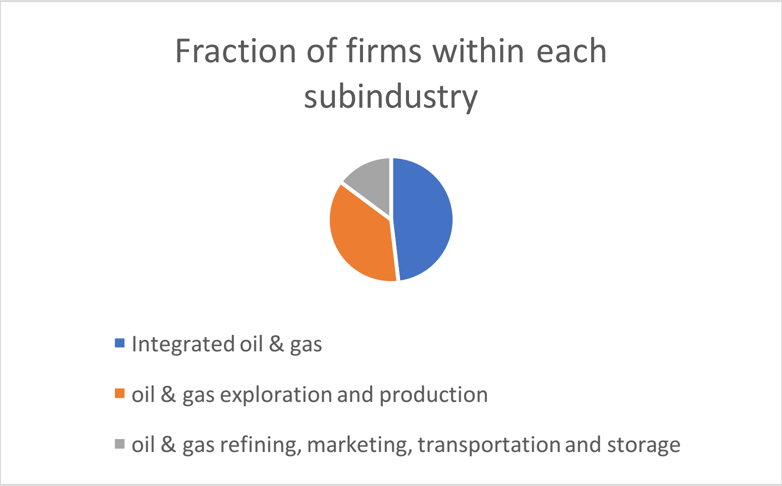

Figure 2 - Fraction of companies in each subindustry. Source: Based on MSCIs' (2020) classification of the companies.

Nearly half of the companies are defined as integrated oil & gas (herby referred to as integrated), which commonly implies that they are involved in the whole value chain of the oil and gas business. Slightly above one-third of the companies belong to the exploration and production sector (referred to as E&P), also known as the upstream segment of the oil and gas industry. These firms typically find, produce, and sell different types of oil and gas (Chen, 2019). The last 15% of the firms are within the refining, marketing, transportation, and storage sector (referred to as RMTS), which is commonly known as the downstream segment of the industry.

3.1.2 Choice of firms

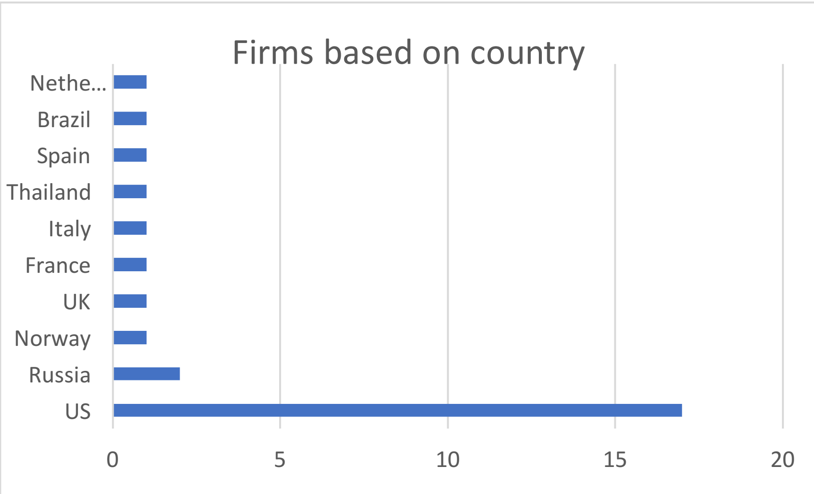

Figure 3 - Frequency of different homecountries for the firms in the research. Source: Based on the classification of home countries presented by MSCI (2020)

To maximize the number of observations in the analysis, as many firms as possible were included. Due to the desire to use updated stock data, the firms had to be listed on a stock exchange.

Another important constraint is that the firms had to be included on the chosen sustainability index and preferably tracked on the index for a more extended period. This necessary condition excludes firms that would be included otherwise and also resulted in a large number of U.S. based firms (illustrated in figure 3). The country factor is considered and controlled for in the analysis.

3.1.3 Choice of sustainability index and agency

An essential task for this research is to identify a suitable ESG-ratings agency. The ESG-ratings agency must include a large number of firms in the industry, and the scores must be tracked a reasonable time backward to have enough data to conduct the analysis. Yahoo finance's ESG-risk ratings (delivered by Sustainalytics) were excluded because many of the firms were not covered, and the ratings do not vary over time.

By using MSCI ESG KLD STATS, Bansal, Wu, and Yaron (2016) managed to expand the time series using data from 1991 to 2011. MSCI ESG KLD STATS only follows U.S firms. It is an index and is, therefore, more suited for SRI research where you can compare the performance of stocks that are included on the index, versus its excluded peers. The data provider is, therefore, not suited for our research, where the emphasis is on the relative ESG performance of the firms.

MSCI's ESG-ratings consider a high number of firms in the industry and follow them over time. Another critical feature of MSCI's ESG-score is that the rating is given based on the company's exposure to ESG-risks and how well the company manages those risks relative to its peers (MSCI, 2020).

The latter is especially important for this paper, due to its focus on capturing differences between the companies in the industry, not across industries. If ESG-scores were given relative to the market as a whole, all firms in this research would probably have poor ESG-scores due to their industry. For these reasons, MSCI is the chosen ESG-rating agency for this analysis.

MSCI's ratings range from AAA (best possible score within the industry) down to CCC (worst score). The ratings are converted into dummy variables to fit into the analysis. With dummy variables, one event is assigned with the value one, and the other event is assigned with the value zero (Wooldridge, 2013, p.182-183). To include a dummy for all the different ESG-scores is problematic because some of the ratings are only assigned to a limited number of firms, while other ratings are more frequent. For example, the rating AAA is only assigned to Equinor, and consequently, a dummy for AAA is then actually a dummy for Equinor itself.

Therefore, the ratings are divided into broader dummy categories, where the number of data per group (dummy variable) increases drastically. The categories are;
Good: AAA, AA, A
Average: BBB, BB
Poor: B, CCC
This categorization is in line with MSCI's classification, only that they refer to such categories as Leader, Average and Laggard, and they include the rating A under the category Average (MSCI, 2020). However, in the analysis, MSCI's classifications (inclusion of rating A under Average) is also tested, and differences are commented.

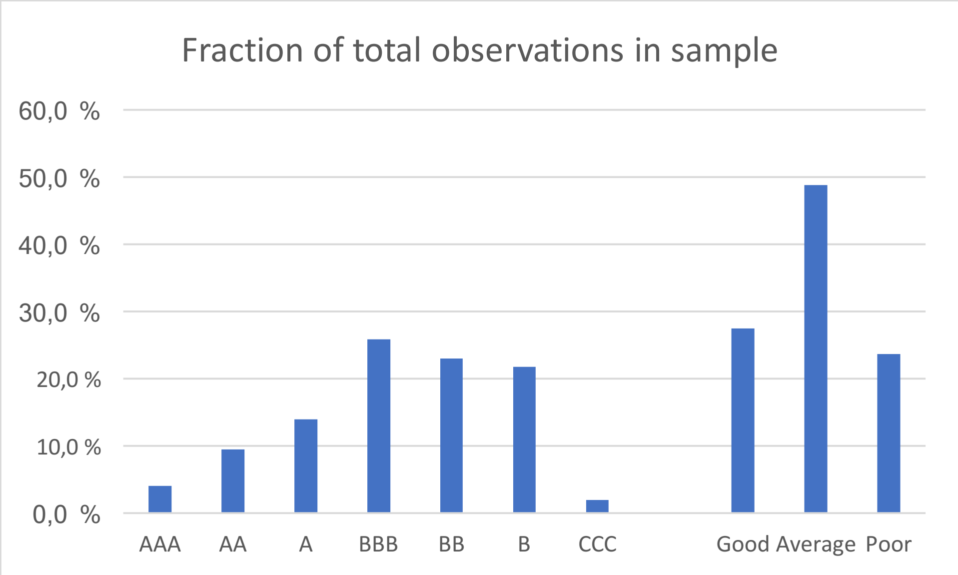

Figure 4 – The sample distribution of ratings, both original ratings from MSCI, and the broader categories used in this research. Source: Based on the ratings and classification presented by MSCI (2020).

Figure 4 shows the fraction of the total observations for the different ratings and the fraction for the combined categories. Including the rating A under category "good" gives an approximately even data distribution between good and poor (around 25% each). Average rating includes the remaining (almost) 50% of observations in the sample.

3.1.4 Factors controlled for in the analysis

This research seeks to identify as many factors as possible, in addition to ESG-rating, that could affect stock return and volatility. Both macroeconomic and financial factors, as well as several dummy variables, are therefore included in the analysis.

Macroeconomic factors

Although the firms in this research (combined) probably have an impact on the oil price through the supply side of the market, this analysis treats it as a given macroeconomic factor in which the firms do not control, as this will be realistic on an individual firm basis. This research has chosen the Brent Crude oil price as a benchmark because it is the most widely used marker of all benchmarks (Kurt, 2020). Any general reference to "the oil price" in this text is hence a reference to the Brent Crude oil price.

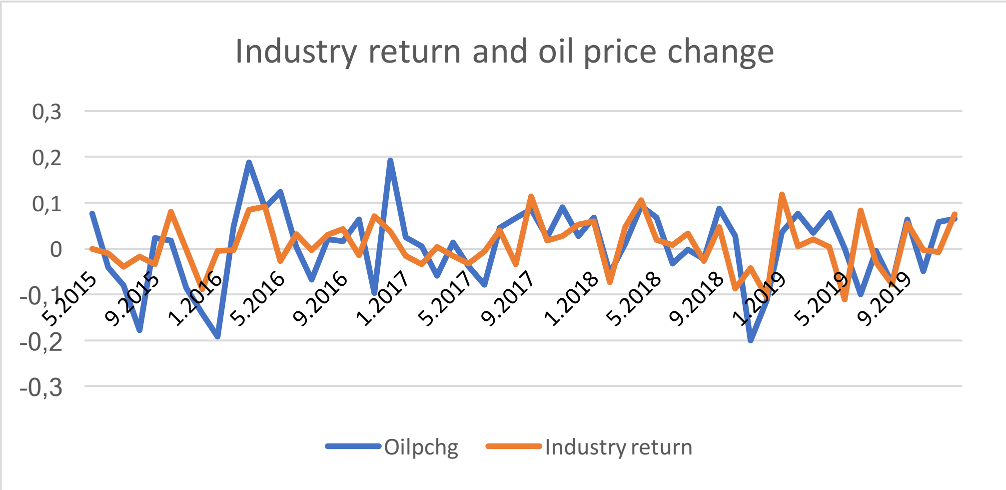

Figure 5 - industry return and oil price change. Source: Based on data from Yahoo Finance (2020).

Figure 5 shows the monthly industry return and the monthly change in the oil price. The monthly industry return is found by adding the return for all individual firms in this study for each month. The correlation-coefficient between industry return and oil price change is calculated to 0,43. Change in oil price is, therefore, included in the analysis of stock returns.

Figure 6 shows the monthly industry volatility, together with the monthly oil price volatility. Monthly industry volatility is found by adding the individual firm's volatility with each other for each month. The factors covary, and it is, therefore, reasonable to include oil price volatility as a control variable in the analysis of stock volatility.

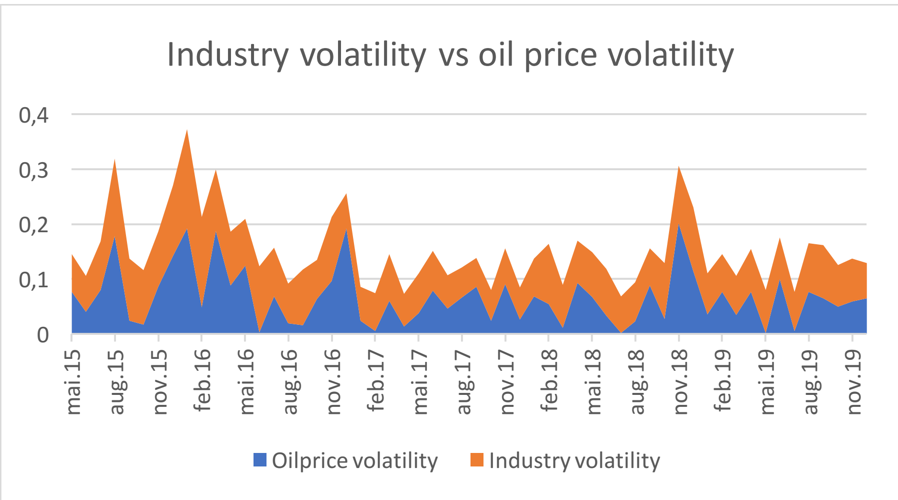

Figure 6 - Industry volatility and oil price volatility. Source: Based on data from Yahoo Finance (2020).

The 10-year Treasury bond yield is included in the analysis as a benchmark for interest rates. Increased interest rates is assumed to offer investors higher risk-free returns, and make stocks relatively less attractive. Hence, increased (decreased) interest rates are expected to have a negative (positive) impact on stock returns and are, therefore, included in the analysis.

Both the oil price and the interest rate are generally assumed to control for broad economic trends in this research. These macroeconomic factors are expected to increase in economic booms as the demand for both energy and capital typically increases when the general activity level in the economy is high, and typically decrease when the activity level is low.

Financial variables

Several financial variables are included in the analysis to control for differences among the firms. Revenues are included as a measure of firm size to control for variations due to size and the stage in a typical business life cycle.

The debt – ratio controls for differences among the firms in terms of gearing effects in the share value. The fraction of earnings before interest, tax, depreciation, and amortization of total revenues (hereby referred to as EBITDA-margin) controls for the impact profitability in operations has on share value and share value volatility.

Dummy variables

The macroeconomic and financial variables are quantitative factors. Binary variables, also called dummy variables, lets us include qualitative control variables (Wooldridge, 2013, p.182-183). In the analysis of stock return and stock volatility, the following dummy variables are included (in addition to the categorized ESG performance dummies described in 3.1.3);

A country dummy variable controls for country-specific factors such as regulations. Figure 3 showed that the home country for most of the firms was the U.S., while all other countries (besides Russia) only was represented with one firm. Individual dummy variables for all countries are, therefore, pointless since this would simply represent the firms themselves for most of the countries. The dummy variable, therefore, takes on the value one if the home country of the firm is the U.S., and zero otherwise, which then controls for differences between the U.S. firms and the firms with another home country. This is important because of the high number of firms from the U.S.

Dummy variables based on the subindustries defined in figure 2 control for differences between them. As their operations differ, it is considered likely that their stock return and risk also might vary. For example, the E&P subindustry is expected to have higher stock volatility than RMTS, as they face the risk of not finding oil and gas.

Dummy variables for the event of upgraded or downgraded ESG-score (referred to as ESG momentum) let us test the possible effect on stock return and volatility of such changes. The variable will take on the value one for the period change occurs, otherwise zero.

Finally, dummies for all years let us control for possible unobserved heterogeneity caused by time effects (Finseraas & Kotsadam, 2013).

3.1.5 Other collection of data

The stock price return and volatility data were found from stock price information downloaded from Yahoo Finance (2020). Adjusted close price adjusts for dividends and splits. Other financial data gathered from this data provider includes interest rates, EBITDA, and debt ratios. Brent Crude oil prices are downloaded from the U.S. Energy Information Administration (2020).

Volatility is measured as the daily standard deviation within the month, found from the deviation of the daily returns from their monthly mean. The daily standard deviation then multiplies with $\sqrt{21}$, which is the assumed average monthly trading days. The procedure is similar to finding the annualized volatility, but for monthly volatility instead of annual.

Of other calculations, the debt ratio is measured as the book value fraction of debt of total assets. The EBITDA margin is calculated as EBITDA fraction of revenues.

3.2 Panel data method and models

This subchapter first briefly present the advantages of the chosen econometrics method for this research, panel data. The panel data models of this research are then presented, as well as the assumptions needed to use these models. Finally, the necessary tests to check for the relative suitability of the models are covered.

3.2.1 Advantages of panel data

Panel data methods are chosen for this research to test the effect of ESG performance on stock return and stock volatility. Panel data increases the number of observations as the same firms are observed over multiple periods of time. The time dimension is what separates panel data from a pooled cross-section and makes it a combination of cross-sectional data and time-series data (Wooldridge, 2013, p.360-361). This feature is ideal for this research as it lets us observe a relatively large number of firms over a considerable period, which has the advantage of increased total observations.

Another advantage of panel data models is that they describe the variation in the dependent variable and the explanatory variables both across firms (between variation), within the firms over time (within variation) and overall (Katchova, 2013). This feature provides a broader understanding of the variables and is described further in chapter 4.

3.2.2 The different panel data models

In this research, three of the most common panel data models are used. These are the pooled model, the random effect model (herby referred to as the RE model), and the fixed effects model (herby referred to as the FE model).

3.2.2.1 Pooled model

The first model is the pooled model, as described (including the equation) by Katchova (2013).

$$y_{it} = \alpha + x_{it}\beta + u_{it}$$

Where y_{it} is the predicted variable for individual i in period t, α is the intercept, β is the coefficient for the explanatory variable x_{it}, and u_{it} is the error term. The pooled model uses constant intercepts, α, and, therefore, does not vary for different firms (Katchova, 2013).

Constant intercepts are not optimal for the data in this research as we want to allow for unobserved effects between the firms. Examples in this research could be effects related to the culture or specific abilities of a particular firm. Such factors are not captured by the pooled model, which thus ignores one of the most significant advantages of panel data (Katchova, 2013).

3.2.2.2 Random-effects model

In the RE model, as described (including the equation) by Katchova (2013), α_i, is included in the error term, and the slope parameters, β, are the same for all individuals. The error term is composited of α_i and e_{it}. The RE model with one explanatory variable is shown below.

$$y_{it} = x_{it}\beta + (\alpha_i + e_{it})$$

The model assumes unobserved heterogeneity between individuals which is captured by α_i. α_i does not change over time but differs between individuals (the i subscript let the intercept vary across individuals). These are unobserved factors across individuals that affect the dependent variable (Katchova, 2013).

With this model, α_i and x_{it} are assumed to be uncorrelated in all periods (Wooldridge, 2013, p. 395). For the dataset in this research, this is not an optimal assumption as we would expect some of the explanatory variables, for example, ESG-performance to be correlated with unobserved firm-specific factors, as firm culture or specific abilities.

3.2.2.3 Fixed effects model

The last presented model is the FE model, as described (including the equation) by Katchova (2013). In this model, the unobserved individual-specific effects α_i can be correlated with the explanatory variables, x_{it}. As shown in the equation, the individual-specific effects, α_i, are included in the model as intercepts. Each firm has a unique intercept, but the slope parameters are equal for all individuals. The fixed-effects model with one explanatory variable is presented below.

$$y_{it} = \alpha_i + x_{it}\beta + u_{it}$$

Because the individual-specific effects and the explanatory variables are allowed to be correlated, explanatory variables that do not vary over time are excluded (Wooldridge, 2013, p. 388). Therefore, explanatory variables such as the home countries or subsectors of the firms are omitted from this model. Problematic assumptions with the pooled model and RE model have been pointed out in the previous sections. Still, one of the reasons to include these models is because they allow for such time-invariant control variables.

3.2.3 Key assumptions

This section describes the key assumptions needed to have unbiased, efficient, and consistent models. These assumptions are tested for the estimated models in chapter 5 (analysis). As the pooled model is not considered particularly important for this research, its assumptions will not be discussed. The key assumptions for the FE model will, therefore, be presented. These assumptions are equivalent to the assumptions of the RE model, although the RE model has the additional assumption that the unobserved effect, α_i, is uncorrelated with all explanatory variables (Wooldridge, 2013, p. 395).

The first assumptions are similar to the assumptions of multiple regression, including that the sample of firms should be random, and there should be no perfect linear relationships between the explanatory variables (Wooldridge, 2013, p.93). These assumptions are assumed to hold, as Stata (the econometrics software) rules out variables with perfect collinearity. Explanatory variables that do not vary over time are automatically omitted from the models. The sample is considered random, but subject to the constraints described in 3.1.2.

However, the most critical assumption to ensure unbiased estimates is the strict exogeneity assumption. This assumption states that the expected value of the error term, u_{it}, should not be correlated with any of the explanatory variables at any point in time (Wooldridge, 2013, p. 388).

The last necessary assumptions are that the error term, u_{it}, are not serial correlated across time, and that the variance of the error term is homoscedastic (Wooldridge, 2013, p. 389).

3.2.4 Tests of suitability

The relative suitability between the different models is checked through a series of tests. The first is the F-test, which automatically follows with the FE model in Stata (including hypotheses).

$H0: all\ u_i = 0$

$H1: all\ u_i \neq 0$

The null hypothesis states that the observed and unobserved (individual-specific) effects are zero. As described in 3.2.2.1, the pooled model does not allow for individual-specific effects. Rejection of the $H0$ will thus implicate that the pooled model is not suitable for the data.

The second test is the Breusch-Pagan Lagrange Multiplier test (hereby referred to as the BP LM-test) presented by Breusch and Pagan (1979).

$H0: Var(u_i) = 0$

$H1: Var\ (u_i)\ \neq 0$

The null hypothesis states that the variance of the observed and unobserved (individual-specific) random effects is zero. Similar to the F-test, a rejection of the null hypothesis implicates that the pooled model is not suitable, and there are significant random effects in the data. If the null hypothesis is rejected, heteroscedasticity might be a problem, and actions to solve the problem should be identified.

If both the null hypothesis of the F-test and the BP LM-test have been rejected, a specification test, called the Hausman-test, first proposed by Hausman (1978) must be conducted. The RE model requires that the unobserved effect, α_i, is uncorrelated with all explanatory variables, which the FE model does not need. The test, therefore, check if the differences in the models' coefficients are systematic or not.

$H0: Differences\ in\ coefficients\ not\ systematic$

$H1: Differences\ in\ coefficients\ systematic$

A rejection of the null hypothesis implies that the additional assumption of the RE model does not hold. In this case, the fixed effects model should be used. If the null hypothesis cannot be rejected, there are low differences in the two sets of coefficients, and both models can be used as the coefficients are similar (Wooldridge, 2013, p.399).

4. Data and statistics

This chapter presents the basic statistics for the different variables that are chosen and collected as input to the panel data set in this research. The features of the different variables are explored and explained before the same variables go into the actual analysis in chapter 5.

This research is concerned with the effect on two different variables, namely weekly stock returns and monthly stock volatility, and is, therefore, divided into two separate parts. First, this chapter presents the statistics for dataset (1) with stock returns as the dependent variable, before the statistics for dataset (2) with stock volatility as the dependent variable is presented.

Some of the variables are given simplified names in the tables and outputs of chapters 4 and 5. Therefore, a table specifying all presented variables can be found in appendix 1.

4.1 The effect of ESG performance on stock returns - statistics

For dataset (1), the dependent variable is weekly stock returns. The other variables are interest rates, dummies for subsector, dummies ESG momentum, country dummy, dummies related to ESG performance, revenues in billions of US $, EBITDA margin, debt ratio, and the change in the oil price. Year dummies are not included as their summary statistics are of limited interest to this research. Statistics for lagged versions of the ESG momentum dummies are also excluded from the statistics review below. However, the statistics of year dummies and lagged ESG momentum can be found in appendix 2.

Table 1 summarizes the statistics and characteristics of the described variables.

Summary statistics, Dataset (1)						
Variable		**mean**	**St.dev.**	**min**	**max**	**observations**
Stock return	overall	0,0017	0,0425	-0,2953	0,3699	N=5919
	between		0,0025	-0,002	0,008	n=27
	within		0,0425	-0,2937	0,3726	T-bar=219,22
E&P dummy	overall	0,3704	0,4829	0	1	N=6615
	between		0,4921	0	1	n=27
	within		0	0,3704	0,3704	T=245
Integrated dummy	overall	0,4815	0,4997	0	1	N=6615
	between		0,5092	0	1	n=27
	within		0	0,4815	0,4815	T=245
RMTS dummy	overall	0,1481	0,3553	0	1	N=6615

	between		0,362	0	1	n=27
	within		0	0,1481	0,1481	T=245
Positive momentum	overall	0,0194	0,1378	0	1	N=6615
	between		0,0272	0	0,102	n=27
	within		0,1351	-0,0827	0,9989	T=245
Negative momentum	overall	0,0145	0,1196	0	1	N=6615
	between		0,0228	0	0,049	n=27
	within		0,1174	-0,0345	0,9655	T=245
Country dummy	overall	0,6296	0,4829	0	1	N=6615
	between		0,4921	0	1	n=27
	within		0	0,6296	0,6296	T=245
Good dummy	overall	0,2747	0,4464	0	1	N=5922
	between		0,3988	0	1	n=27
	within		0,1949	-0,4559	1,2358	T-bar=219,33
Average dummy	overall	0,488	0,4999	0	1	N=5924
	between		0,425	0	1	n=27
	within		0,2692	-0,473	1,353	T-bar=219,41
Poor dummy	overall	0,2369	0,4252	0	1	N=5926
	between		0,3893	0	1	n=27
	within		0,1857	-0,6281	0,6717	T-bar=219,48
EBITDA margin	overall	0,2925	0,2259	-0,4283	0,865	N=5173
	between		0,1899	0,0629	0,7731	n=26
	within		0,1229	-0,2225	0,6596	T-bar=198,96
Debt ratio	overall	0,5154	0,1027	0,249	0,753	N=5646
	between		0,0995	0,3102	0,6772	n=27
	within		0,0318	0,4489	0,6499	T-bar=209,11
Revenues	overall	70,8969	87,635	1	389	N=6038
	between		85,3602	1,6408	310,7143	n=26
	within		18,1581	-5,8173	149,1827	T-bar=232,23
Oil price change	overall	0,0013	0,0407	-0,1489	0,1437	N=6588
	between		0	0,0013	0,0013	n=27
	within		0,0407	-0,1489	0,1437	T=244
Interest rates	overall	0,0229	0,0044	0,0137	0,0323	N=6615
	between		0	0,0229	0,0229	n=27
	within		0,0044	0,0137	0,0323	T=245

Table 1 - summary statistics and characteristics of variables in dataset (1) with stock return as the dependent variable.

The panel dataset consists of 245 weeks (t), for 27 different companies (n), which makes the dataset consist of 6615 observations (N) for each variable. Some omitted data for the different variables is expected because, for example, stock returns are measured from the time MSCI started giving ESG-scores to the different firms (which occurred at different periods for different companies). However, most firms have data over the whole period (t), and the dataset is therefore defined as strongly balanced by the analysis in Stata.

The table shows the different overall, within (across time factor) and between (across firms' factor) statistics for each variable, which is an important feature of panel datasets. The overall variation consists of differences over time and across firms. Variables with zero within variation are classified as time-invariant factors, while variables with zero between variation are classified as firm-invariant factors (Katchova, 2013).

Oil price change and interest rates are examples of firm-invariant factors, meaning that they do not differ across firms. The country dummy and the dummies for subsectors are examples of time-invariant factors, which means they do not vary over time, which is logical since a firm's home country or subsector typically does not change. The debt ratio has low within variation (SD of 0,0318) while the between variation is much higher (SD of 0,0995), hence the variable is close to time-invariant, or at least do not change much over time. Low within variation for the debt ratio suggests that the industry as a whole has a quite stable book value of debt to assets over the data period.

The weekly stock return has an average of 0,17% for the industry. The standard deviation within each firm over time is 4,25%, which is a lot larger than the standard deviation between the firms 0,25%. The time factor within the firms is of more importance to describe the variation in returns, than factors between the firms. This implies that it can be difficult to find firm-specific factors that explain differences in stock returns across firms, for example, related to ESG performance.

For ESG performance categories, we observe that most of the variation is between the categories compared to within. This is expected because the score for each firm is likely to be correlated over time, while independent between firms (Katchova, 2013).

The companies in the paper differ significantly on different parameters. In terms of size, the smallest companies have revenues of 1b$ while the largest saw 389b$ in the best year. The

EBITDA margin is ranging from -42,83% to 86,50%. It is important to control for such large differences by including the factors as regressors in the analysis.

ESG momentum

Table 1 shows that most variation for the positive momentum is individual invariant within the firms (SD of 0,1351). As the dummy takes on the value 1 for the different firms when MSCI changes a firm's ESG-score, the variation will be more substantial across time than across individuals. However, some variation is found across firms (SD of 0,0272) as well because some firm's ESG-score is consistent over the whole period, while other's change.

The same pattern is clear for the negative momentum. Here, the standard deviation across time (within) is 0,1174, and the standard deviation across firms (between) is 0,0228.

Table 1 shows ESG momentum is represented with 1-2% (subject to the chosen length of period which is discussed in chapter 5) of the observations, hence up- and downgrades are relatively rare events. The number of upgraded and downgraded ESG-scores each year is shown in figure 7.

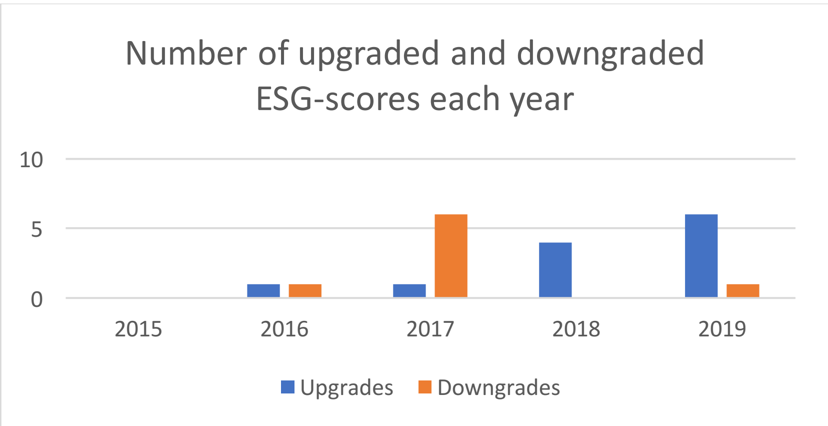

Figure 7 - Number of upgraded and downgraded ESG-scores each year. Source: Based on rating data from MSCI (2020).

In 2015 there were no changes, and in 2016 only two firms received a new ESG-score. From the year 2017, more frequent changes in the scores occur. This lets us test the effect of momentum in general. However, because of the low number of observations at the start of the period (2015 and 2016), it is difficult to test if investors show an increasingly stronger interest in ESG over time.

4.2 The effect of ESG-score on stock volatility – statistics

This part considers the statistics for dataset (2), where the dependent variable is stock volatility. All variables from dataset (1) are included in this dataset as well, but only the statistics for variables that are not included in dataset (1) are shown here. The focus is, therefore, on stock volatility and oil price volatility.

Table 2 summarizes the statistics and some of the features of these variables.

Summary statistics, dataset (2)						
Variable		**mean**	**st.dev.**	**min**	**max**	**observations**
Stock volatility	overall	0,0857	0,0433	0,0141	0,3446	N=1511
	between		0,0246	0,0514	0,1357	n=27
	within		0,0359	0,0079	0,3301	T-bar=55,96
Oil price volatility	overall	0,0665	0,0505	0,0013	0,2009	N=1512
	between		0,0000	0,0665	0,0665	n=27
	within		0,0505	0,0013	0,2009	T=56

Table 2 - summary statistics and characteristics of variables in dataset (2) with stock volatility as the dependent variable. All other variables from dataset (1) are also included in the dataset (2), but not presented in the table.

This panel dataset consists of 56 time periods (months) and (still) 27 firms, which gives a total of 1512 observations per variable (for variables with no missing data). The dataset is described as unbalanced. This is not considered problematic as the cause for the missing data is because (for example) returns are measured from the time MSCI started to provide ESG-scores to the different firms (which happened at different periods in time).

The variable stock volatility has a mean of 0,0857, and the variation is distributed between firms and within the specific firm across time. The most substantial part of the variation is across time, with a standard deviation of 0,0359. Still, a large portion of the variation is between the firms, with a standard deviation of 0,0246. In the previous subchapter (4.1), we saw that the between variation of stock return (the other dependent variable) was very low, which makes it challenging to explain the differences across firms later in the analysis. In comparison, the portion of between variation for volatility is much higher, which suggests it should be easier to describe the differences in volatility across firms than stock returns.

The oil price volatility is the same for all firms, so there is no between variation since the variable is firm invariant. The average oil price volatility is 0,0665, and the standard deviation, which is equal to the variation across time (within factor), is 0,0505.

5. Analysis

This chapter presents the results from the analysis of the data material. Both datasets are described in terms of statistical results and the different variables coefficients. The structure of the presentation is similar to the previous chapter, where dataset (1) with stock returns are presented first before dataset (2) with stock volatility is presented. Further discussions of the results will take place in chapter 6.

For both datasets, the F-test, BP LM-test, and the Hausman-test (when needed) are presented. For dataset (2), these tests are presented more briefly as they are conducted in the same manner. The purpose of the tests is to check the suitability of the different models.

All coefficients are examined for statistical significance at the 1% ***, 5% **, and 10%* level. Hence, a higher number of * means lower p-value and higher statistical significance of the coefficient.

For each model, two models are presented because dummies for ESG performance are assumed to be affected by ESG momentum. Therefore, these variables should not be run within the same model as this could dilute some of the effects ESG performance (and vice versa) potentially have on stock returns and stock volatility. The purpose is to reduce the problem of endogeneity, and further explanations are presented in chapter 7.

Only the models (a) with dummies for ESG performance (and all other explanatory variables, excluding ESG momentum) are presented in complete form. From the models (b) with ESG momentum, only the results for the ESG momentum are presented, with no additional control variables or general properties of the models. This is because other variables and the properties do not change much between these models. However, the complete models (b) with ESG momentum are found in appendix 3,4,5 and 6.

Lagged dummies for ESG momentum are also included in the models (b) for 1, 2, and 3 periods after the change occurs. These let us explore when the effects (if any) occur, and for how long the effect is evident. Periods of one month has been tested, but resulted in a meager number of data for each dummy variable and was therefore not chosen. Periods of a quarter of a year solved this problem.

Both the homogeneity and the no autocorrelation assumptions are graphically checked for violations. When violations are found, measures are taken to reduce the problem. The strict exogeneity condition is assumed to hold because of the actions taken against the problem, including the separation of assumed correlated explanatory variables, and the inclusion of a high number of control variables. However, such an assumption is bold, and, therefore, problematized further in chapter 7.

5.1 The effect of ESG performance on stock returns - analysis

The first analysis is conducted on dataset (1). As described in the previous section, models (b) with dummies for ESG momentum (including lags) are run separately from the models (a) with dummies for ESG performance.

5.1.1 F-test

The start of the analysis is to run a FE model with the chosen variables. In Stata, the F-test follows automatically. The null and alternative hypothesis are;

$H0: all\ u_i = 0$

$H1: all\ u_i \neq 0$

The results give an F(25, 4945) of 0,50. With a p-value of 0,9815, the null hypothesis cannot be rejected. Thus, it cannot be rejected that the observed and unobserved (individual-specific) effects are zero, and as a consequence, the pooled model cannot be rejected. Both the FE model and the pooled model are, therefore, presented.

5.1.2 Breusch-Pagan Lagrange Multiplier-test

For a similar purpose as the F-test, the BP LM-test tests the variance of the observed and unobserved random effects, with the following hypotheses;

$H0: Var(u_i) = 0$

$H1: Var\ (u_i) \neq 0$

The results from the test gives a Chibar2(01) of 0,00. With a p-value of 1,000, the null hypothesis cannot be rejected at any level of significance. As the test finds no evidence against the null hypothesis, the RE model is inappropriate for this analysis.

The failure to reject the null hypothesis at any level of significance, suggests the residuals exhibit little or none heteroscedasticity, which is further confirmed by a graphical check of the residuals plots in Stata. Hence, it can be assumed that the model exhibits little heteroscedasticity.

5.1.3 Hausman-test

Failure to reject the null hypothesis at any level of significance in the BP LM-test means that the RE model is not appropriate for this dataset. It is, therefore, not necessary to carry out the Hausman-test to decide whether to use the RE model or the FE model.

5.1.4 Fixed effects-model (Dataset 1, model 1A and 1B)

This section presents the results from the FE model (1A and 1B) on dataset (1) with stock return as the dependent variable. A visual check of the plot of the residuals showed clear patterns of positive serial correlation, which violates the no autocorrelation assumption. Autocorrelation might give standard errors that are too small, and hence, lower p-values. Robust standard errors address this problem and provide the results presented below.

FE model 1A, dataset (1)		
Number of obs =		4982
Number of groups =		26
Obs per group:	min =	135
	avg =	191,6
	max =	210
F(11, 25) =		13,60
Prob > F =		0,000
R-sq:		
	within =	0,0132
	between =	0,0096
	overall =	0,0086
corr(u_i, Xb) =		-0,5244

Table 3 - Summary of results from the fixed effects model 1A in dataset (1) with stock return as dependent variable.

Stock return	Coefficient	Robust SE	t	p>|t|	(95% Conf. Intevall)	
Interest rates	0,0021	0,0017	1,25	0,224	-0,0014	0,0057
Integrate dummy	0,0000	(omitted)				
E&P dummy	0,0000	(omitted)				
RMTS dummy	0,0000	(omitted)				
Year 2015 dummy	-0,0651***	0,0128	-5,09	0,000	-0,0915	-0,0388
Year 2016 dummy	0,0084***	0,0016	5,31	0,000	0,0052	0,0117
Year 2017 dummy	0,0009	0,0013	0,69	0,497	-0,0018	0,0036
Year 2018 dummy	-0,0019	0,002	-0,95	0,351	-0,006	0,0022
Year 2019 dummy	0,0000	(omitted)				
Country dummy	0,0000	(omitted)				
Good dummy	-0,0014	0,0034	-0,42	0,676	-0,0085	0,0056
Average dummy	0,0014	0,0019	0,74	0,464	-0,0025	0,0052
Poor dummy	0,0000	(omitted)				
EBITDA margin	0,0005	0,0039	0,14	0,889	-0,0074	0,0085
Debt ratio	0,0117	0,019	0,62	0,543	-0,0274	0,0509
Revenue	0,00003	0,00002	1,54	0,136	-0,00001	0,00008
Oil price change	0,0189	0,0181	1,05	0,306	-0,0184	0,0563
Constant	-0,0132	0,0102	-1,29	0,21	-0,0342	0,0079
Sigma_u	0,0041					
sigma_e	0,0412					
rho	0,0099					

Table 4 - Summary of results from the fixed effects model 1A in dataset (1).

Stock return	Coefficient	Robust SE	t	p>|t|	(95% Conf. Intevall)	
Upgrade t	0,0065	0,0045	1,45	0,159	-0,0027	0,0157
Upgrade t+1	-0,0021	0,0026	-0,79	0,434	-0,0075	0,0033
Upgrade t+2	-0,0007	0,0038	-0,18	0,86	-0,0084	0,0071
Upgrade t+3	-0,0055**	0,0022	-2,48	0,02	-0,0101	-0,0009
Downgrade t	0,0019	0,0031	0,60	0,552	-0,0045	0,0082
Downgrade t+1	0,0059*	0,0033	1,78	0,088	-0,0009	0,0128
Downgrade t+2	0,0043	0,0043	1,61	0,121	-0,0012	0,0098
Downgrade t+3	-0,0011	0,004	-0,27	0,786	-0,0093	0,0071

Table 5 - Summary of results from the fixed-effects model 1B in dataset (1) with momentum as explanatory variable.

Tables 3 and 4 summarize the results from model 1A. Model 1B with dummies for momentum (including lags) are presented in table 5 (whole model 1B is found in Appendix 3).

Table 3 shows the F-statistics and the corresponding p-value of 0,000. Model 1A is, therefore, statistically significant at all levels, and the null hypothesis that the R-squared is zero is rejected. The model is not explaining much of the variation in stock returns with within, between, and overall R-squared of 0,0132, 0,0096, and 0,0086, respectively.

The dummies for subsectors and the country dummy are examples of time-invariant variables. For each individual (firm), the FE model takes the value at each period and subtracts the mean of the whole period. For time-invariant variables, the value at each period is the same, and also equals the mean. The value will be zero, and the variable is not included in the FE model (Katchova, 2013). The dummies for subsectors and the country dummy are omitted because of this.

Poor dummy and year 2019 dummy are omitted because of collinearity. These dummies are part of "dummy groups" of several variables. By omitting one of the variables, the other can be viewed in comparison to the omitted variable. Hence, good and average ESG performance dummies are measured against poor ESG performance dummy.

The only statistically significant coefficients at the 1% level are the year dummy 2015 and year dummy 2016, which both have p-values of 0,000. Table 5 shows that the upgrade dummy in period t+3 has a p-value of 0,02, and the downgrade dummy in period t+1 has a p-value of 0,0828. These lagged ESG momentum dummies are hence significant at the 5% and 10% level of significance, respectively.

Upgrade in period t+3 has a negative coefficient, which implies reduced stock returns in period 3 after a firm has received an upgraded ESG-score. On the contrary, downgrade in period t+1 has a positive coefficient, which means stock returns are predicted to increase in period 1 after a downgraded ESG-score. Both results may seem contradictory to expectations and will be discussed further in chapter 6.

Significance at the chosen levels is not found for the financial variables EBITDA-margin, debt ratio, and revenue. A likely reason is that expectations about future financials are more important for the investors' valuation of the stock than their historical records. A bigger surprise is the failure to find a significant relationship between the stock price and the oil price changes. In chapter 3.1.4, the oil price and the combined industry average return was calculated to have a correlation coefficient of 0,43. However, the model is not considering the average return, but the individual firms return. A consequence is that potential outliers (of the observations) will not be average out, as they would have been with the industry average. This again leads to an increased magnitude of outliers, which might explain the lack of significance.

The dummies for ESG performance fails to reject the null hypothesis. P-values for the good and average dummy is 0,676 and 0,464, respectively. It cannot, therefore, be rejected that these coefficients are zero. ESG performance within the oil and gas industry, therefore, has no explanatory power in this model.

Usage of MSCI's classification of ESG performance categories (inclusion of the rating A under the category Average), was also tested but provided no major differences in results compared to this research's classification.

The rho of 0,0099 shows the percentage (approximately 1%) of the variation that is explained by individual-specific or unobserved effects. This meager number means that most of the variation is explained by idiosyncratic effects, which is the observation specific error term (Katchova, 2013, 9:51). The low variation explained by the individual-specific effects, is a likely reason for the failure to reject the pooled model for this dataset. This is because the pooled model assumes constant coefficients, which is not a very problematic assumption when the individual-specific effects are at such low levels.

5.1.5 Pooled model (dataset 1, model 2A and 2B)

Both the F-test and BP LM-test failed to reject the pooled model in favor of the FE model and the RE model. Therefore, the pooled model is presented. Because the individual-specific effects did not explain much of the variation the pooled model's assumption of constant coefficients is not very problematic. However, as there is still some individual-specific effects in the dataset, the primary model for this dataset is the FE model, and the pooled model is presented more briefly.

Pooled model 2A, dataset (1)	
Number of obs =	4982
F(14, 4967) =	4,11
Prob > F =	0,000
R-squared =	0,0139
Root MSE =	0,0411

Table 6 - Summary of results from the pooled model 2A in dataset (1) with stock return as dependent variable.

Stock return	Coefficient	Robust SE	t	p>|t|	(95% Conf. Interval)	
Interest rates	0,0020	0,0022	0,92	0,358	-0,0023	0,0064
Integrate dummy	-0,0023	0,0020	-1,16	0,244	-0,0063	0,0016
E&P dummy	-0,0014	0,0020	-0,71	0,480	-0,0052	0,0025
RMTS dummy	0,0000	(omitted)				
Year 2015 dummy	-0,0642***	0,0130	-4,92	0,000	-0,0898	-0,0386
Year 2016 dummy	0,0093***	0,0032	2,93	0,003	0,0031	0,0155
Year 2017 dummy	0,0021	0,0020	1,03	0,305	-0,0019	0,0059
Year 2018 dummy	0,0000	(omitted)				
Year 2019 dummy	0,0020	0,0025	0,78	0,433	-0,0029	0,0068
Country dummy	-0,0037**	0,0017	-2,15	0,031	-0,0070	-0,0003
Good dummy	-0,0009	0,0021	-0,46	0,649	-0,0052	0,0032
Average dummy	-0,0004	0,0017	-0,23	0,821	-0,0036	0,0029
Poor dummy	0,0000	(omitted)				
EBITDA margin	-0,0002	0,0040	-0,06	0,951	-0,0081	0,0077
Debtratio	-0,0026	0,0070	-0,37	0,714	-0,0164	0,0112
Revenue	0,00000	0,00000	-0,32	0,745	-0,00002	0,00001
Oil price change	0,0189	0,0178	1,06	0,289	-0,0160	0,0537
Constant	0,0008	0,0079	0,10	0,923	-0,0148	0,0163

Table 7 - Summary of results from the pooled model 2A in dataset (1).

Stock return	Coefficient	Robust SE	t	p>|t|	(95% Conf. Interval)	
Upgrade t	0,0061	0,0037	1,63	0,104	-0,0012	0,0134
Upgrade t+1	-0,0023	0,0033	-0,7	0,483	-0,0088	0,0042
Upgrade t+2	-0,0010	0,0043	-0,24	0,810	-0,0094	0,0073
Upgrade t+3	-0,0062	0,0044	-1,41	0,158	-0,0147	0,0024
Downgrade t	0,0008	0,0047	0,17	0,861	-0,0084	0,0102
Downgrade t+1	0,0049	0,0040	1,23	0,217	-0,0029	0,0127
Downgrade t+2	0,0036	0,0044	0,81	0,420	-0,0051	0,0123
Downgrade t+3	-0,0019	0,0046	-0,4	0,689	-0,0109	0,0072

Table 8 - Summary of results from the pooled model 2B in dataset (1) with momentum as explanatory variable.

Tables 6 and 7 show the output of the pooled regression model 2A. The F-statistics is 4,11, and the corresponding p-value is 0,000. The overall model is, therefore, statistically significant at all levels, and the null hypothesis that the R-squared is 0 is rejected. The estimated R-squared is 0,0139, and the model, therefore, has a minimal impact on explaining the variation in the weekly stock return. Table 8 shows the results from model 2B, which is with ESG momentum instead of ESG performance.

RMTS dummy, the year 2018 dummy, and the poor dummy are automatically omitted from the model due to collinearity.

Most of the variables in the model are not significant at any adequate level of significance. The only significant variables are year dummy 2015, year dummy 2016, and the country dummy, with p-values of 0,000, 0,003, and 0,031, respectively. The country dummy has a negative coefficient, which implies that firms with the U.S. as home country, has significantly lower returns than firms with other home countries.

The dummy variables for ESG performance, good and average, have p-values of 0,649 and 0,821, respectively. Hence, in this model as well, the null hypothesis that these coefficients are different from zero cannot be rejected.

Table 5 shows that none of the ESG momentum dummies are statistically significant at the chosen levels, in contrast to the same variables in the FE model. Their p-values range from 0,104 to 0,861, and, therefore, it cannot be rejected that the coefficients are different from zero.

5.2 The effect of ESG performance on stock volatility- analysis

This last analysis is conducted on dataset (2) with monthly volatility as the dependent variable and interest rates, oil price volatility, EBITDA margin, debt ratio, revenue, dummies for subsectors, year dummies, country dummy, dummies for ESG momentum (including lags) and dummies for ESG performance category as the explanatory variables. Similar to dataset (1), models (b) with dummies for ESG momentum (including lags) are separated from the models (a) with dummies for ESG performance.

The models are tested for violations of the necessary assumptions, as with the previous dataset. If violations are found, measures are taken towards the problems to minimize the bias in the results.

5.2.1 F-test, BP LM-test and Hausman-test

Because the methodology is similar to chapter 5.1, the hypothesis-tests will not be defined here. The start of the analysis is to run the FE model with the chosen variables. With an F-test statistic of 14,24 and a p-value of 0,000, the observed and unobserved (individual-specific) effects are significantly different from zero. Therefore, the pooled model is not presented as it does not allow for individual-specific effects. The FE model is more suitable and is presented because it allows for these effects.

The BP LM-test of the variance of the random effects has a chibar2(01) of 624,78, and a p-value of 0,000. It can, therefore, be rejected that the variance of the error term is zero, at all levels of significance. There are significantly random effects in the data, and the RE model is, therefore, a more suitable model than the pooled model.

Rejecting that the variance of the error term is zero suggests the residuals could be heteroscedastic. Visual checks of the residual plot in Stata confirms this. To address this problem, the dependent variable (stock volatility) is log-transformed. Natural log-transformation is possible since the variable has only positive values. The same is done for the control variable revenue. After the transformation, the problem of heteroscedasticity was severely mitigated.

Both the F-test and the BP LM-test revealed that the pooled model was not suited for the data. The Hausman specification test should, therefore, be carried out to test the relative suitability of the RE model and the FE model. The RE model requires that the unobserved effect, α_i, is uncorrelated with all explanatory variables, which is not needed for the FE model. It is, therefore, tested if the differences in the models' coefficients are systematic or not.

$H0: Differences\ in\ coefficients\ not\ systematic$

$H1: Differences\ in\ coefficients\ systematic$

The test has a chi2(10) of 9,11. With a p-value of 0,5220, we fail to reject the null hypothesis that the difference in the coefficients is not systematic. Both models can be presented as the coefficient do not differ much.

5.2.2 Fixed effects model

The natural log-transformed FE model (dataset 2, model 1A) was first run without any further adjustments since we assumed that both homoscedasticity (after log-transformation) and strict exogeneity (although assumption criticized in chapter 7) holds. However, a visual check of the plot of the residuals showed clear patterns of a positive correlation between them. The no autocorrelation assumption is therefore violated. Robust standard errors address this problem, as with the analysis in 5.1, which gives the results presented in tables 9, 10, and 11 below.

Fixed effects model 1A, dataset 2		
Number of obs =		1128
Number of groups =		26
Obs per group:	min =	30
	avg =	43,4
	max =	48
F(10, 1092) =		33,82
Prob > F =		0,000
R-sq:		
	within =	0,2671
	between =	0,4356
	overall =	0,0022
corr(u_i, Xb) =		-0,623

Table 9 - Summary of results from the fixed-effects model 1A in dataset (2) with stock volatility as the dependent variable.

Log(stock volatility)	Coefficient	Robust SE	t	p>|t|	(95% Conf. Intevall)	
Interest rates	-0,1588***	0,0355	-4,48	0,000	-0,2318	-0,0858
Integrate dummy	0,0000	(omitted)				
E&P dummy	0,0000	(omitted)				
RMTS dummy	0,0000	(omitted)				
Country dummy	0,0000	(omitted)				
Good dummy	-0,2535**	0,1204	-2,11	0,045	-0,5014	-0,0056
Average dummy	-0,1347	0,0904	-1,49	0,149	-0,3208	0,0514
Poor dummy	0,0000	(omitted)				
EBITDA margin	-0,1582*	0,0864	-1,83	0,079	-0,3361	0,0197
Debt ratio	0,5793	0,3593	1,61	0,12	-0,1608	1,3193
Log(revenue)	0,0976	0,0906	1,08	0,291	-0,0889	0,2842
Year 2015 dummy	0,0000	(omitted)				

Year 2016 dummy	0,1625**	0,0759	2,14	0,042	0,0063	0,3187
Year 2017 dummy	-0,1761***	0,0462	-3,81	0,001	-0,2713	-0,0808
Year 2018 dummy	0,1887***	0,0416	4,53	0,000	0,1030	0,2745
Year 2019 dummy	0,0000	(omitted)				
Oil price volatility	0,9204***	0,1641	5,61	0,000	0,5824	1,2583
Constant	-2,7626***	0,3523	-7,84	0,000	-3,4882	-2,037
Sigma_u	0,4381					
sigma_e	0,2975					
rho	0,6844					

Table 10 - Summary of results from the fixed effects model 1A in dataset (2).

Log(stock volatility)	Coefficient	Robust SE	t	p>│t│	(95% Conf. Interval)	
Upgrade t	-0,1137**	0,0532	-2,14	0,042	-0,2232	-0,0042
Upgrade t+1	-0,0932	0,0861	-1,08	0,29	-0,2705	0,0841
Upgrade t+2	0,0658	0,0609	1,08	0,29	-0,0597	0,1913
Upgrade t+3	-0,0457	0,0729	-0,63	0,537	-0,1959	0,1045
Downgrade t	-0,0407	0,0799	-0,51	0,615	-0,2054	0,1239
Downgrade t+1	-0,0419	0,0592	-0,71	0,485	-0,1639	0,0799
Downgrade t+2	-0,107***	0,0388	-2,77	0,01	-0,1872	-0,0275
Downgrade t+3	0,0305	0,0481	0,63	0,532	-0,0685	0,1295

Table 11 - Summary of results from the fixed-effects model 1B in dataset (2) with momentum as explanatory variable.

Table 9 show the F(10, 1092) of 33,82 with a p-value of 0,000. The overall model is statistically significant at all levels of significance, and the null hypothesis that the R-squared of the model is 0 is rejected. The model, therefore, explains some of the variations in the monthly volatility. The model explains 26,71% of the variation in monthly volatility within each firm and 43,56% between the firms. The overall R-squared is 0,22%.

As for the FE model described in 5.1.3, time-invariant variables, including both the dummies for subsector and country dummy, are omitted. Poor dummy, the year 2015 dummy, and year 2019 dummy are omitted due to collinearity.

Interest rates (0,0000), year 2017 dummy (0,001), year 2018 dummy (0,000), oil price volatility (0,000), downgrade in period t+2 (0,01) and the constant (0,000) are statistically significant at the 1% level, with their respective p-values in parenthesis. Good ESG performance dummy (0,045), year 2016 dummy (0,042) and upgrade in period t (0,042) are statistically significant at the 5% level, and EBITDA-margin (0,079) is statistically significant at the 10% level.

Interest rates have a negative coefficient, which is in line with expectations. Higher interest rates will typically make stocks less attractive relative to bonds, and lead to a lower stock return, and less stock trading. Less trading means decreased stock volatility. Among the other significant control variables, the oil price volatility has, as expected, based on figure 6 in chapter 3.1.4, a positive coefficient. Large shocks in the oil price are hence predicted to increase the firms' stock volatility.

ESG performance and changed ESG-score

The dummy for poor ESG performance is omitted, and therefore the good and average ESG performance is considered in comparison to the poor ESG performance dummy.

The dummy for good ESG performance is statistically significant at the 5% level and has a negative coefficient of -0,2535. Because the dependent variable is log-transformed, the relative difference in monthly stock volatility between good and poor ESG performance is approximately 22,4% ($100 * (e^{-0,2535}$ -1)). Hence, monthly volatility is estimated to drop by more than one fifth when going from a poor to a good ESG performance, controlling for other factors.

The coefficient for the average ESG performance is negative at -0,1347. However, with a p-value of 0,149, the coefficient is not statistically significant at any of the chosen levels.

Using MSCI's classification of categories (inclusion of rating A under average) had little effect on good and average ESG performance. Average ESG performance was still not statistically significant, while the coefficient for good ESG performance was slightly decreased, and had a minor increase in the p-value.

The upgrade in period t is statistically significant, with a p-value of 0,042. The coefficient is negative at -0,1137. The downgrade in period t+2 has a significant (p-value of 0,01) negative coefficient of -0,107. Hence, controlling for other factors, volatility is predicted to drop with approximately one-tenth relative to firms with unchanged ESG scores, for firms who receive an upgraded (in period zero) or a downgraded (in period two) ESG score. A drop in the volatility following both upgrade and downgrade is surprising and will be further discussed in chapter 6.2.3.

A rho of 0,6844 means individual-specific effects explains 68,44% of the variation in the model. The variation is distributed with approximately two-thirds of individual-specific effects and one-third idiosyncratic effects (Katchova, 2013, 9:51). Compared to the FE model for dataset (1), we observe that in this model, the percentage of variation explained by the individual-specific effects is a lot greater.

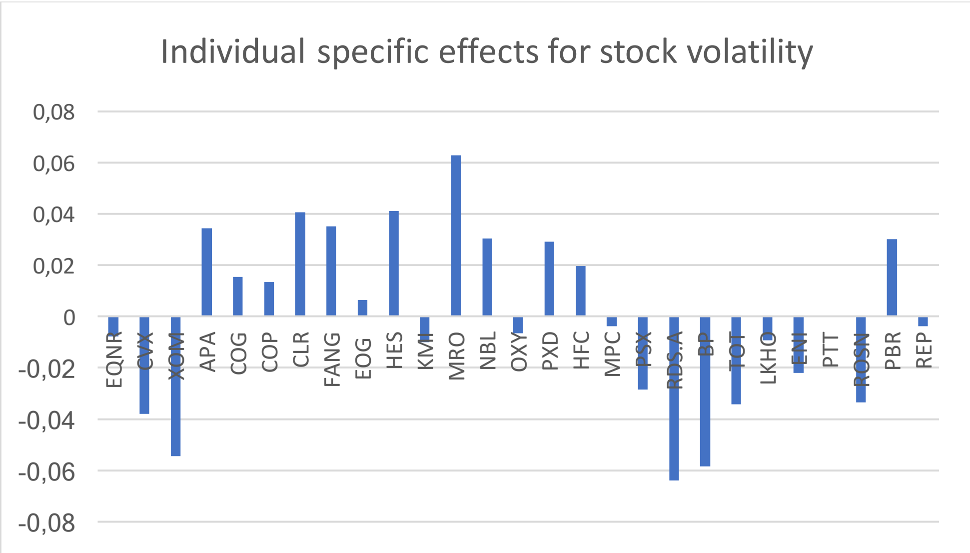

Figure 8 - Distribution of estimated individual-specific effects (for each individual firm, identified by thicker, on x-axis) on stock volatility (on y-axix). Source: Estimations done in Stata.

As described in subchapter 3.2.2.3 the FE model allows for a unique intercept for each individual in the sample. Figure 8 shows the estimated individual (firm) specific effects, $\hat{\alpha}_i$, which is the estimated intercept for each firm with the FE model. Deviations from zero means there are firm-specific effects that make the firm have different monthly stock volatility than what the model predicts.

5.2.3 Random-effects model

Tables 12 and 13 summarizes the results from the RE model (2A) on dataset (2) with monthly volatility as the dependent variable. Table 14 summarize the results from the RE model (2B) with ESG momentum as an explanatory variable instead of ESG performance.

The RE model has an additional assumption compared to the FE model. Table 9 showed a correlation-coefficient of -0,623 between the unobserved effects and the explanatory variables, and the assumption is, therefore, violated. This makes the RE model less appropriate for this dataset. Therefore, the FE model is considered the primary model. In contrast, the RE model is presented because it lets us explore the effect of time-invariant variables, and allows us to compare potential differences in the models.

Random effects model 2A, dataset 2		
Number of obs =		1128
Number of groups =		26
Obs per group:	min =	30
	avg =	43,4
	max =	48
Wald chi2(18) =		512,74
Prob > chi2 =		0,0000
R-sq:		
	within =	0,2645
	between =	0,6344
	overall =	0,4372
corr(u_i, X) =		0 (assumed)

Table 12 - Summary of results from the random-effects model 2A in dataset (2) with stock volatility as dependent variable.

Log (stock volatility)	Coefficient	Robust SE	z	p>|z|	(95% Conf. Interval)	
Interest rates	-0,1586***	0,0354	-4,49	0,000	-0,2279	-0,0893
Integrate dummy	-0,2148*	0,1188	-1,81	0,071	-0,4476	0,0181
E&P dummy	0,3007**	0,1526	1,97	0,049	0,0017	0,5997
RMTS dummy	0,0000	(omitted)				
Country dummy	-0,1759	0,1079	-1,63	0,103	-0,3875	0,0357
Good dummy	-0,2453**	0,11	-2,23	0,026	-0,4609	-0,0296
Average dummy	-0,1378	0,0865	-1,59	0,111	-0,3072	0,0317
Poor dummy	0,0000	(omitted)				
EBITDA margin	-0,1358*	0,0816	-1,66	0,096	-0,2958	0,0242
Debt ratio	0,6649**	0,3013	2,21	0,027	0,0743	1,2554
Log(revenue)	-0,0337	0,0468	-0,72	0,471	-0,1254	0,0579
Year 2015 dummy	0,0000	(omitted)				
Year 2016 dummy	0,1023	0,0628	1,63	0,104	-0,0209	0,2254
Year 2017 dummy	-0,2064***	0,041	-5,03	0,000	-0,2868	-0,126
Year 2018 dummy	0,1916***	0,0427	4,48	0,000	0,1078	0,2753
Year 2019 dummy	0,0000	(omitted)				
Oil price volatility	0,9185***	0,1644	5,59	0,000	0,5963	1,2408
Constant	-2,2699***	0,3393	-6,69	0,000	-2,935	-1,6049
Sigma_u	0,1935					
sigma_e	0,2975					
rho	0,2972					

Table 13 - Summary of results from the random-effects model 2A in dataset (2).

Log(volatility)	Coefficient	Robust SE	z	p>|z|	(95% Conf. Interval)	
Upgrade t	-0,1197**	0,0509	-2,35	0,019	-0,2195	-0,0199
Upgrade t+1	-0,0994	0,0825	-1,2	0,228	-0,2611	0,0623
Upgrade t+2	0,0514	0,0643	0,8	0,424	-0,0747	0,1774
Upgrade t+3	-0,0637	0,0649	-0,98	0,327	-0,1909	0,0636
Downgrade t	-0,0427	0,0823	-0,52	0,604	-0,2040	0,1186
Downgrade t+1	-0,0460	0,0624	-0,74	0,460	-0,1683	0,0762
Downgrade t+2	-0,1122***	0,0398	-2,82	0,005	-0,1902	-0,0342
Downgrade t+3	0,0294	0,0474	0,62	0,535	-0,0636	0,1224

Table 14 - Summary of results from the random-effects model 2B in dataset (2) with momentum as explanatory variable.

The overall model is statistically significant at all levels with a wald chi2(18) of 512,74 and a corresponding p-value of 0,000. The model explains 26,45% of the variation in volatility within the different firms and 63,44% between the firms. Overall the explanatory variables in the model explain 43,72% of the variation in volatility. RMTS dummy, poor dummy, the year 2015 dummy, and year 2019 dummy are omitted due to collinearity.

In general, the coefficients of the variables (both direction and size), and their p-values, give similar results with the RE model as with the FE model. However, the dummies for the subsectors integrated and E&P was omitted from the FE model because they are time-invariant variables. Both variables are statistically significant, with p-values of 0,071 and 0,049, respectively. In line with expectations, firms within the E&P subsector has increased stock volatility, compared to firms within the RMTS subsector, while firms within the integrated subsector have decreased volatility, controlled for all other factors.

ESG performance and ESG momentum

The dummy for poor ESG performance is omitted, and the good and average dummy is, therefore, considered in comparison to the poor dummy.

The good dummy, which is significant at the 5% level, has a negative coefficient of -0,2453. As the dependent variable is log-transformed, the relative difference in monthly stock volatility between good and poor ESG performance is approximately 21,75% ($100 * (e^{-0,2453}-1)$). Hence, volatility is estimated to drop by slightly more than one fifth when going from a poor to a good ESG performance, all else held equal.

The average dummy is not significant at the levels of statistical significance chosen for this research, with a p-value of 0,111. The coefficient is negative at -0,1378, but because of the lack of statistical significance, the results cannot be generalized. Both the good and average ESG performance dummy has relatively similar results in the RE model, as with the FE model.

Similarly, the statistically significant dummies for ESG momentum, are the same as for the FE model. These are the coefficients for the upgrade in period t, and downgrade in period t+2 are statistically significant (on the 5% and 1% level, respectively). As in the FE model both have negative coefficients.

6. Discussion of results

This chapter discusses the most relevant results from the analysis in chapter 5. The discussion takes both previous research and theory into account to explain the results. The most critical implications the results could have for investors are also discussed. Both analyses in the previous chapter had its primary focus on the FE model. The results from this model are the primary source for these discussions, although the results from the other models also provide insights.

6.1 The effect of ESG performance on stock returns

From dataset (1), two models were presented. Both the FE model and the pooled model was overall significant but did not explain much of the variation in stock returns. In general, the emphasis of valuations is on future expectations rather than historical records. Because the model does not capture future expectations directly, this may explain the lack of explanatory power of the model.

In chapter 4.1, we identified and pointed out that most of the variation in stock returns are across time within each firm (4,25%), and minimal variation across firms (0,25%). The implication of this is that there is a very limited variation across firms left for the model to explain. Therefore, it is not a surprise that most of the explanatory variables fail to explain the differences across firms.

Both models were not able to conclude, at any level of statistical significance, that there was any difference in weekly stock returns between the good, average, and poor ESG performers in the industry. The statistical insignificance implicates that there is no clear relationship between ESG-score and the stock return of the firms in the industry.

6.1.1 Possible explanations

Findings of no clear relationship correspond with previous studies lack of consensus regarding the effect of ESG performance on stock returns. An insufficient amount of data is a common explanations for the failure to establish a clear (statistically significant) relationship between ESG performance and stock return. In studies where relationships are identified, the results often contradicts each other. For example, Bansal, Wu, and Yaron (2016) finds a positive effect of responsible behavior, while Marsat and Williams (2011), finds the exact opposite. This can, to a large extent, be explained by differences in periods, choice of sustainability index, or other issues related to the methodology of the research.

A possible reason to the lack of findings in this research is that investors might be indifferent to ESG performance within the industry. Investors' attention to sustainability has been well documented over the past few years. However, the focus is often broader; for example, they would prefer firms that produce renewable energy sources rather than non-renewable. Perhaps the investors that care most about sustainability would not invest in the oil and gas industry at all, and therefore not care about the relative performance of the firms within the industry. If this is the case, then investors that decide to invest in the oil and gas industry would be indifferent or at least have less prominent preferences in terms of ESG performance. Then other factors than ESG performance would describe the differences in stock returns between the firms in the industry.

Another possible explanation is that the investors perform a careful examination of the different firms and their ESG performance, but that the benefits and costs of good, average and poor ESG performance, with the assessment procedure proposed by Weber (2008), do not differ. If this is the case, then the ESG effort (or lack off) does not increase (decrease) value. With this view, investors care about ESG performance within the industry, but the stock return of the different ESG performers should not differ based on ESG-score because the benefits and costs equal to zero.

6.1.2 Possible implications

No clear relationship between ESG performance and stock return does not necessarily imply that ESG is of no interest. The stakeholder and shareholder positions were presented in chapter 2.3, where the shareholder position argues that the sole purpose of the company is to maximize shareholder wealth. Proponents of this view would invest in the poor ESG performing companies if this provided higher risk-adjusted (risk will be discussed in chapter 6.2) stock returns than the good ESG performing companies. This research's lack of findings should make these investors indifferent between the firms.

Assuming the only reason to invest in the poor ESG performing firms is that they provide excess risk-adjusted returns compared to the good ESG performing firms, then the indifference is effectively an argument for the good ESG performing companies. This is because there is no loss of wealth associated with these investments compared to the poor ESG performing firms. With this view, the burden of proof is on the poor ESG performers, who need to show higher stock returns than the good performers to attract investors.

However, an obvious pitfall with this argument is that it cannot necessarily be claimed that wealth optimizing investors will invest in the good ESG performers when they are indifferent between the investment opportunities. Hence, no clear results regarding the effect of ESG performance on stock returns imply that these investors might invest in all firms, no firms, or identify other valuation criteria or methods, as described below.

Another pitfall with the argument is that it relies on historical information about the stock returns. If investors who seek maximization of their wealth, think the stocks of the good ESG performing firms are overvalued, they will expect the corresponding stock prices to drop, making these investments suboptimal. The core is that the individuals' beliefs about the future have implications on their preferences, and therefore may affect the investment decision.

As discussed, the implications of no clear relationship between ESG performance and stock returns ultimately depends on the investors' perception. However, the essential reason why stock returns are not very interesting independently is because of the dependent relationship between risk and return, which is discussed in subchapter 6.2.

6.1.3 The effect of ESG momentum on stock returns – possible explanations

The effect of upgraded and downgraded ESG-score, also called momentum, was included in the analysis (although in separate models) of stock returns. This subchapter discusses the results and their possible explanations. The periods after changed ESG-score was set to 12 weeks, to increase the data for each dummy variable. Periods of 4 weeks were dismissed due to the minimal number of data associated with such small periods.

ESG momentum was not statistically significant for the first periods (t) after a change, in neither of the models. In the FE model, for the lagged periods, significant coefficients for the upgrade in period t+3 (negative coefficient and p-value of 0,02) and downgrade in period t+1 (positive coefficient and p-value of 0,088) was identified. In contrast, none of the lagged periods was significant for the pooled model. Inconsistency between the models and the relatively high p-value of the downgrade in t+1 contributes to doubt regarding the results.

Because of the failure to identify any relationship between ESG performance in general and stock returns, it is not a surprise that the findings regarding the relationship between changed ESG-score and stock returns are inconclusive. If changed ESG-score affected stock returns, this would suggest that there should be an effect on stock returns from general ESG performance, at least in the long run. This likely relationship is, in fact, the reason why the explanatory variables were run separately in the analysis.

The most likely explanation that a conclusive relationship between changed ESG-score and stock return cannot be identified is due to the lack of long periods of data. Giese and Lee (2019) also point to the need for extended time series to conclude with certainty regarding momentum, although Giese and Nagy (2018) found weak evidence of upgraded (downgraded) ESG scores leading to higher (lower) stock returns. Such results contradicts the findings of this research.

Altogether, the effects of changed ESG-score on stock return is weak and contradictive to previous research. The results are, therefore, considered inconclusive. As previous research points out, more extended data periods are needed to conclude with certainty.

6.2 The effect of ESG performance on stock volatility

From the analysis of the effect ESG performance had on stock volatility, the FE model and the RE model was presented. Both models were overall statistically significant. The FE model explains 26,71% of the variation within each firm and 43,56% of the variation between the firms. The overall R-squared is 0,22%. Similarly, the RE model explains 26,45% of the variation within the firms and 63,44% of the variation between the firms. Overall the explanatory variables in this model explain 43,72% of the variation in monthly stock volatility.

Both models hence explain a relatively large portion of the variation in volatility, and has explanatory power, especially between the firms. For both models, the analysis finds that the good ESG performers have significantly lower monthly stock volatility than the poor ESG performers, with reduced volatility of approximately one-fifth. None of the models find significant results between the average and poor ESG performers.

6.2.1. Possible explanations

The most robust explanation to the findings of lower volatility for firms with good ESG performance is that these firms are less likely to be involved with scandals and receive negative press publicity and attention, which again leads to lower stock volatility. This explanation has support in research where Giese and Lee (2019) find that companies with low (poor) MSCI ESG rating historically have a higher frequency of significant ESG related incidents. Additionally, by conducting research on Australian stocks, Smales (2014) finds that negative non-scheduled news has a high impact on trading volume and consecutively on volatility. Similar results on non-scheduled news are found at the London Stock Exchange by Groß-Klußmann and Hautsch (2011). Hence, poor ESG performers generally receive more negative attention, and these studies find that volatility, as a consequence, will be higher for these firms.

The core of the MSCI ESG rating process is to identify the individual firms' exposure to ESG risk and how well they manage those risks (MSCI, 2020). It is, therefore, expected that firms with more ESG related risks also should have higher volatility, that is, higher risk. The findings of this research serve as a support to the validity and efficiency of MSCIs rating system. It also confirms that their assessment of risks is relevant and a useful tool to examine the riskiness of firms.

Diversified sources of revenue are another possible explanation for the lower volatility of good ESG performers. Firms who increase their investments in renewable energy sources are likely to strengthen their ESG-scores. Receiving revenues from multiple sources of energy will reduce the exposure to specific energy prices because of diversification, as long as the prices are not perfectly positively correlated. Lower exposure to shocks in specific energy commodity prices should, in turn, lead to lower stock volatility.

6.2.2 Possible implications

In contrast to the analysis for stock returns, a statistically significant relationship between ESG performance and stock volatility has successfully been identified. In subchapter 6.1, it was argued that investors with the ultimate target of wealth maximization, would be indifferent between the firms in this sample because of the lack of difference in stock returns between good, average and poor ESG performers. However, this was without taking risk into account.

The relationship between risk and return is one of the most fundamental principles of finance, where you should be rewarded proportional to the levels of risk you take (Ackert & Deaves, 2010, p.8). When an investor considers stocks that do not differ on expected return, she will choose the stock that has the relatively lowest levels of risk. In the context of the stakeholder and shareholder positions, this means that even the proponents of the shareholder position will invest in the good ESG performers as this will maximize value, subject to risk levels.

However, finance generally considers stocks in terms of their attributions to a broad portfolio of stocks and commonly assumes fully diversified investors. If investors are fully diversified, the additional risk of poor ESG performers will disappear, making the poor performers equally attractive. Then, one could question whether it matters with increased risk for the poor performers?

The answer is that it should depend on two factors. First of all, if the additional risk is systematic or diversifiable, and secondly, if investors are adequately diversified. If both of the conditional factors are satisfied, the extra volatility of the poor ESG performers relative to the good ESG performers should not matter to investors.

Giese and Lee (2019) find that companies with good ESG ratings generally are more adaptable to changing market conditions and environments, which they refer to as lower systematic risk. This argues that the additional risk of the poor ESG performers is not diversifiable. Concerning investors' level of diversification, Statman (2004) finds that the optimal level of stocks needed to exploit the total benefits of diversification entirely is over 300, while the average investors hold 3 or 4 stocks. Besides, French and Poterba (1991), find that most investors hold nearly all their stocks in domestic markets. Hence, they do not exploit the possibility of international diversification either.

The extra risk of poor ESG performers are thus not diversifiable, and investors, in general, are not fully exploiting the benefits of diversification. Therefore, with the findings of this research, the most rational investor behavior is to invest in the firms with good ESG performance, since the additional risk of poor performers cannot be adequately diversified away. This recommendation is valid for all investors, also those who do not have any preferences regarding sustainability, as it will maximize the wealth of the investor.

There are, however, two conceptual situations where investors should not follow this recommendation. First, in the unlikely event that the investor has abnormal risk preferences, and prefers risk, then she should not follow the advice, as it is based on risk-averse preferences. The second situation occurs when investors have reason to believe historical data are not representative of the future. Under such circumstances, it could be more rational to invest in the poor ESG performers if this is expected to offer a higher risk-adjusted return.

6.2.3 The effect of ESG momentum on stock volatility – possible explanations

The effect of ESG momentum was included in the analysis (in separate models) of stock volatility. This subchapter discusses the results and possible explanations for the findings.

The analysis of momentums effect on stock volatility, had statistically significant results for the upgrade in period t and downgrade in period t+2, in both the FE model and the RE model. In the FE model, their p-values were 0,042 and 0,01, respectively, and both coefficients were negative. Both the firms who receive an upgrade or a downgrade is predicted to decrease the monthly volatility with approximately one-tenth, all else held equal, in period 0 and 2 respectively.

The findings are more robust than the results of ESG momentum's effect on stock return because the findings have lower p-values and are found in both models. The immediate decrease in stock volatility following an upgrade could have similar explanations as the general lower stock volatility of good ESG performers, which includes a lower risk of controversies and (or) lower exposure to specific energy commodity prices. However, if this is the case, the effect would be expected to last in all periods, not only the first.

Possible explanations for the lagged decrease in stock volatility following a downgraded ESG-score are hard to find in both previous research, theory, and other parts of this research. The most likely reason is that the limited data on the downgraded ESG-scores leads to results that are caused by other (omitted and unknown) variables.

Altogether, the effects of changed ESG-score on stock volatility, as with stock returns, seems rather inconclusive. As previous research points out, more extended data periods are needed to conclude with certainty.

7. A critical view, limitations, and further research

During the work with this research, a lot of choices are made, and restraints have been taken. There are lots of pitfalls, dilemmas, and trade-offs, which are presented in this section.

A critical view and limitations

The first issue is related to causality and the necessary assumptions needed to ensure unbiased models. To identify that a variable has a causal relationship on another variable is quite different than identifying a general relationship between the variables. In an attempt to identify causal effects, strategies presented by Finseraas and Kotsadam (2013) have been used to eliminate as much as possible of the potential sources of errors. Such measures are first that highly correlated explanatory variables are omitted from the model (automatically by Stata). Secondly, time dummies are included in the model to reduce the problem of unobserved heterogeneity caused by effects over time. Lastly, variables related to ESG performance are run in models separately from variables related to changes in ESG-score, as the latter is expected to affect the ESG performance. Hence, explanatory variables that are expected to be correlated are separated.

The strict exogeneity condition was assumed to hold for the models. However, it is unrealistic that we have controlled for all factors that could have an effect on the dependent variable, even though macroeconomic factors, the most important financial variables and a large set of dummy variables are included in the models. It is also unlikely that the independent variables do not affect each other at all, even though separate models have been run in an attempt to avoid this for the most important explanatory variables.

Because of this, I have chosen to define the statistically significant results as predictions based on the model throughout the text, to point out that the relationships are not necessarily strictly causal. It should, however, be pointed out that the problem with strict exogeneity is evident in nearly all research, and that the main objective of this research is to reduce the sources of errors to a minimum.

Other issues that should be addressed are those regarding ESG-scores. First of all, the general use of ESG-scores as a tool to identify the effect of focus on sustainability can be questioned. ESG-scores are ultimately the product of a combination of qualitative and quantitative data and are, therefore, subject to evaluations that are not strictly objective. As a consequence, the ratings may be biased, and the relative differences among firms may not represent the real difference in terms of sustainability. Further, the choice of the specific agency provider for the ESG-scores in this research, MSCI, limits the analysis to the explanatory power of these ratings specifically.

Transformation into broader ESG performance categories was chosen to address the problem of possible subjectivity and bias in the ratings. Transformation into broader categories should reduce the effect of possible biased ratings because more data are collected per variable. A negative consequence is that it limits the research to these self-defined categories. The categorization within each category is also subjective and might affect the results. The categorization used by MSCI was also tested in the analysis to address deviations from different categorizations. However, the results with the different categorizations were, in large, without notable differences.

Further research

This research is limited to the explanatory power of the ratings from MSCI. Potential further research can build on the findings of this research, and test if the same effect on stock volatility (and lack of effect on stock returns) is found by using other rating agencies. Other agencies would likely contain ESG information on different firms than those in this research. Therefore, by using another rating agency, one could test the effect on stock volatility with a different sample of firms. Along with the possibility of generalizing the results across different rating agencies, such research would also provide insights into the relative performance of the agencies in identifying (ESG related) risks.

The need for more extended periods of data has been pointed out in both previous studies and this research. As time proceeds and the dataset becomes longer, the relationship between stock returns and ESG performance should be reexamined. Extended periods of data would also result in more changes in ESG-scores, which again will provide a better foundation for research on the effect of ESG momentum.

As most research on the effect of ESG is conducted on a broad market portfolio, increased industry-specific research would help shed light on potential differences in the relative explanatory power of ESG between industries.

8. Conclusion

Investors have shown increased interest in firms that emphasize with the surrounding environment. Understanding the effect such investments have on the investors' risk-adjusted return is analyzed comprehensively during the last few years. This research examines the effect on investors in the oil & gas industry. By doing so, the research provides new insights into a specific industry that stands at a strategic cross-road concerning its future in a world where sustainability is likely to be at the center of attention.

Using ESG-scores as a proxy to measure firms' attention to sustainability, this research considers 27 firms in the industry for 4,5 years, analyzed with panel data methodology.

The analysis finds no significant relationship between sustainability and stock returns. The effects of ESG momentum (changed ESG-score) on both stock return and stock volatility is also inconclusive. However, significant results concerning the impact of sustainability on stock volatility are identified. Firms that have a good ESG-performance experience a decrease in the volatility of about one fifth, compared to the volatility of firms with poor ESG-performance. Investors should, therefore, invest in oil & gas companies with good ESG-performance, to maximize their risk-adjusted return, and hence, their wealth.

Bibliography

Ackert, L. F., & Deaves, R. (2010). *Behavioral Finance - Psychology, Decision-Making, and Markets.* Mason: South-Western Cengage Learning.

Banerjee, N. Song, L. & Hasemyer, D. (2015, Sept 16th). Exxon's Own Research Confirmed Fossil Fuels' Role in Global Warming Decades Ago. Retrieved March 13, 2020, from https://insideclimatenews.org/news/15092015/Exxons-own-research-confirmed-fossil-fuels-role-in-global-warming

Bansal, R., Wu, D., & Yaron, A. (2016). Socially Responsible Investing: Good is Good, Bad is Bad. Retrieved from https://pdfs.semanticscholar.org/a629/8d73f6d6e06f7900f1e352f75e1dc3677989.pdf?_ga=2.76756309.1991255841.1591880503-1368607682.1591880503

Borchgrevink, A. S. (2019). *Giganten* (Vol. 2019). Oslo: Kagge forlag.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica,* 47(5), 1287-1294. https://www-jstor-org.ezproxy.uis.no/stable/1911963

Chen, J. (2019, Apr 21). Exploration & Production (E&P). Retrieved Mar 23, 2020, from https://www.investopedia.com/terms/e/exploration-production-company.asp

Dedekam Jr, A. (2002). *Mikroøkonomi* (2. edition 2007). Bergen: Fagbokforlaget.

Degnes, O. G. (2020, Feb 11). Kjell Inge Røkkes oljeselskap Aker BP utelukker satsing på fornybar energi. *Finansavisen*. Retrieved Feb 12, 2020, from https://finansavisen.no/nyheter/energi/2020/02/11/7496749/kjell-inge-rokkes-oljeselskap-utelukker-satsing-pa-fornybar-energi

Finseraas, H., & Kotsadam, A. (2013). Hvordan identifisere årsakssammenhenger i ikke-eksperimentelle data? En ikke-teknisk introduksjon. *Tidsskrift for samfunnsforskning, 54*(3), 371-387. https://www-idunn-no.ezproxy.uis.no/tfs/2013/03/hvordan_identifisere_aarsakssammenhenger_i_ikke-eksperimente

French, K. R., & Poterba, J. M. (1991). Investor Diversification and International Equity Markets. *The American Economic Review,* 81(2), 222-226. https://www-jstor-org.ezproxy.uis.no/stable/2006858

Giese, G., & Lee, L.-E. (2019). Weighing the evidence: ESG and Equity returns. *Research Insight,* 2019 (April), 1-14, https://www.msci.com/documents/10199/9aec76d8-376f-91ef-a575-b2b0ea65061a

Giese, G., & Nagy, Z. (2018). How markets price ESG - Have changes in ESG score affected stock prices? *Research Insight,* 2018 (November), 1-25. https://www.msci.com/documents/10199/f3f252a7-0dbd-035b-e67c-964a87fbcd03

Groß-Klußmann , A., & Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Emirical Finance,* 18(2), 321-340. https://doi.org/10.1016/j.jempfin.2010.11.009

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251-1271. https://www-jstor-org.ezproxy.uis.no/stable/1913827

Katchova, A. (2013). Panel Data Models. Retrieved from Econometrics Academy Website: https://docs.google.com/file/d/0BwogTI8d6EEiX2ZGeTRObjktOVk/edit

Katchova, A (Producer). (2013). *Panel Data Models in Stata.* [Video clip]. Retrieved from Econometrics Academy Website: https://sites.google.com/site/econometricsacademy/econometrics-models/panel-data-models

Kurt, D. (2020, Jan 4). Benchmark Oils: Brent Crude, WTI and Dubai. Retrieved Feb 12, 2020, from https://www.investopedia.com/articles/investing/102314/understanding-benchmark-oils-brent-blend-wti-and-dubai.asp

Marsat, S., & Williams, B. (2011). CSR and Market Valuation: International Evidence. *International Conference of the French Finance Association (AFFI),* May 11-13,2011,1-24, http://dx.doi.org/10.2139/ssrn.1833581

MSCI. (2020). ESG Ratings. Retrieved Jan 9, 2020, from MSCI ESG Ratings: https://www.msci.com/esg-ratings

Nilsen, K. P., & Noergaard, R. W. (2011). CSR and mainstream investing: a new match? - an analysis of the existing ESG integration methods in theory and practice and the way forward. *Journal of Sustainable Finance & Investment*, 1(3-4), 209-221. https://www-tandfonline-com.ezproxy.uis.no/doi/full/10.1080/20430795.2012.655889

PwC. (2019). *Risikopremien i det norske markedet (November 2019)*. Retrieved from https://www.pwc.no/no/publikasjoner/pwc-risikopremie-2019.pdf

Ritchie, H., & Roser, M. (2017, May). CO2 and Greenhouse Gas Emissions. *Our World in Data.* Retrieved February 10, 2020, from https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions

Smales, L. A. (2014). Non-scheduled news arrival and high-frequency stock market dynamics: Evidence from the Australian Securities Exchange. *Research in International Business and Finance*, 32, 122-138. https://doi.org/10.1016/j.ribaf.2014.03.006

Statham, M. (2004). The Diversification Puzzle. *Financial Analysts Journal, 60*(4), 44-53. https://doi-org.ezproxy.uis.no/10.2469/faj.v60.n4.2636

Statoil. (2001). *The future is now; Statoil and sustainable development* (Sustainability report 2001) . Retrieved from https://www.equinor.com/no/investors/our-dividend/annual-reports-archive.html

Tirole, J. (2006). *The Theory of Corporate Finance*. Princeton: Princeton University Press.

U.S. Energy Information Administration. (2019). *International Energy Outlook 2019 with projections to 2050.* Retrieved from https://www.eia.gov/outlooks/ieo/pdf/ieo2019.pdf

U.S. Energy Information Administration. (2020). Petroleum & Other Liquids. Retrieved from https://www.eia.gov/dnav/pet/hist/rbrteD.htm

Weber, M. (2008). The business case for corporate social responsibility: A company-level measurement approach for CSR. *European Management Journal,* 26(4), 247-261. https://doi.org/10.1016/j.emj.2008.01.006

Wooldridge, J. M. (2013). *Introduction to Econometrics* (Europe, Middle East & Africa Edition ed., Vol. 2014). Hampshire: Cengage Learning EMEA.

Yahoo Finance. (2020). Retrieved from https://finance.yahoo.com/

Zhou, M. (2019, Aug 22). ESG, SRI, and Impact Investing: What's the Difference? Retrieved Apr 1, 2020, from https://www.investopedia.com/financial-advisor/esg-sri-impact-investing-explaining-difference-clients/

Appendices

Appendix 1

Variable name	Spesification
Stock return	Weekly stock return (adjusted for dividends and splits)
Interest rates	The 10-year Treasury bond yield
Integrate dummy	Dummy for the subsector Integrated oil & gas
E&P dummy	Dummy for the subsector Exploration and production
RMTS dummy	Dummy for the subsector Refining, marketing, transportation and storage
Year 2015 dummy	Dummy for the year 2015
Year 2016 dummy	Dummy for the year 2016
Year 2017 dummy	Dummy for the year 2017
Year 2018 dummy	Dummy for the year 2018
Year 2019 dummy	Dummy for the year 2019
Country dummy	Dummy for home country. Takes the value 1 if firm is based in the U.S., zero otherwise
Good dummy	Dummy for the ESG-scores classified as Good
Average dummy	Dummy for the ESG-scores classified as Average
Poor dummy	Dummy for the ESG-scores classified as Poor
EBITDA-margin	Revenue / Earnings before interest, tax, depreciation and amortization
Debt ratio	Book value of debt / Book value of total assets
Revenue	Revenues in billions of USD
Oil price change	Weekly oil price change
Upgrade t	Dummy for the event of upgraded ESG-score (when the change occur)
Upgrade t+1	Dummy for the event of upgraded ESG-score 1 period after change occur
Upgrade t+2	Dummy for the event of upgraded ESG-score 2 periods after change occur
Upgrade t+3	Dummy for the event of upgraded ESG-score 3 periods after change occur
Downgrade t	Dummy for the event of downgraded ESG-score (when the change occur)
Downgrade t+1	Dummy for the event of downgraded ESG-score 1 period after change occur
Downgrade t+2	Dummy for the event of downgraded ESG-score 2 periods after change occur
Downgrade t+3	Dummy for the event of downgraded ESG-score 3 periods after change occur
Log(stock volatility)	Natural logarithm of the monthly stock volatility
Oil price volatility	Monthly volatility in the oil price

Appendix 1 – Specification of all presented variables in this research.

Appendix 2

Summary statistics for Year dummies, changed ESG-score lags						
Variable		**mean**	**St.dev.**	**min**	**max**	**observations**
Upgrade t+1	overall	0,0147	0,1202	0	1	N=6615
	between		0,0265	0	0,0979	n=27
	within		0,1174	-0,0833	1,0106	T=245
Upgrade t+2	overall	0,013	0,1133	0	1	N=6615
	between		0,0224	0	0,0571	n=27
	within		0,1111	-0,0441	0,964	T=245
Upgrade t+3	overall	0,0113	0,1059	0	1	N=6615
	between		0,0206	0	0,049	n=27
	within		0,1039	-0,0376	0,9991	T=245
Downgrade t+1	overall	0,0144	0,119	0	1	N=6615
	between		0,0226	0	0,0531	n=27
	within		0,1169	-0,0387	0,9735	T=245
Downgrade t+2	overall	0,0127	0,1119	0	1	N=6615
	between		0,0219	0	0,0489	n=27
	within		0,1099	-0,0363	0,9637	T=245
Downgrade t+3	overall	0,0127	0,1119	0	1	N=6615
	between		0,0219	0	0,0489	n=27
	within		0,1099	-0,0363	0,9637	T=245
Year 2015 dummy	overall	0,1504	0,3575	0	1	N=6615
	between		0,0031	0,1347	0,151	n=27
	within		0,3575	-0,0006	1,016	T=245
Year 2016 dummy	overall	0,2088	0,4064	0	1	N=6615
	between		0,0031	0,2082	0,2245	n=27
	within		0,4064	-0,0157	1,0006	T=245
Year 2017 dummy	overall	0,2122	0,4089	0	1	N=6615
	between		0	0,2122	0,2122	n=27
	within		0,4089	0	1	T=245
Year 2018 dummy	overall	0,2164	0,4119	0	1	N=6615
	between		0,0008	0,2163	0,2204	n=27
	within		0,4119	-0,0039	1,0002	T=245
Year 2019 dummy	overall	0,2121	0,4088	0	1	N=6615
	between		0,0008	0,2082	0,2122	n=27
	within		0,4088	-0,0002	1,0039	T=245

Appendix 2 – summary statistics for year dummies and lagged momentum dummies, in dataset (1).

Appendix 3

Fixed effects stock return		
Number of obs =		4982
Number of groups =		26
Obs per group:	min =	135
	avg =	191,6
	max =	210
F(17, 25) =		37,72
Prob > F =		0,0000
R-sq:		
	within =	0,0144
	Between =	0,0123
	overall =	0,0107
corr(u_i, Xb) =		-0,4402

Stock return	Coefficient	Robust SE	t	p>∣t∣	(95% Conf. Intevall)	
Interest rates	0,0026	0,0017	1,50	0,146	-0,0010	0,0063
Integrate dummy	0,0000	(omitted)				
E&P dummy	0,0000	(omitted)				
RMTS dummy	0,0000	(omitted)				
Year 2015 dummy	-0,066***	0,0129	-5,07	0,000	-0,0922	-0,0389
Year 2016 dummy	0,008***	0,0017	4,69	0,000	0,0044	0,0115
Year 2017 dummy	0,0001	0,0016	0,08	0,936	-0,0031	0,0033
Year 2018 dummy	-0,0028	0,0022	-1,29	0,21	-0,0072	0,0017
Year 2019 dummy	0,0000	(omitted)				
Country dummy	0,0000	(omitted)				
Upgrade t	0,0065	0,0045	1,45	0,159	-0,0027	0,0157
Upgrade t+1	-0,0021	0,0026	-0,79	0,434	-0,0075	0,0033
Upgrade t+2	-0,0007	0,0038	-0,18	0,86	-0,0084	0,0071
Upgrade t+3	-0,0055**	0,0022	-2,48	0,02	-0,0101	-0,0009
Downgrade t	0,0019	0,0031	0,60	0,552	-0,0045	0,0082
Downgrade t+1	0,0059*	0,0033	1,78	0,088	-0,0009	0,0128
Downgrade t+2	0,0043	0,0043	1,61	0,121	-0,0012	0,0098
Downgrade t+3	-0,0011	0,004	-0,27	0,786	-0,0093	0,0071
EBITDA margin	0,0004	0,0039	0,11	0,911	-0,0074	0,0085
Debt ratio	0,0091	0,0211	0,43	0,67	-0,0344	0,0526
Revenue	0,00003	0,00002	1,62	0,118	-0,00001	0,00008
Oilprice change	0,0169	0,0185	0,92	0,367	-0,0211	0,055
Constant	-0,0122	0,0112	-1,08	0,289	-0,0353	0,0109

Sigma_u	0,0037
sigma_e	0,0411
rho	0,0079

Appendix 3 - Summary of results from the complete fixed-effects model 1B in dataset (1) with momentum as explanatory variable.

Appendix 4

Pooled model stock return	
Number of obs =	4982
F(20, 4967) =	3,21
Prob > F =	0
R-squared =	0,0151
Root MSE =	0,0411

Stock return	Coefficient	Robust SE	t	p>|t|	(95% Conf. Interval)	
Interest rates	0,0026	0,0022	1,16	0,247	-0,0018	0,0070
Integrate dummy	-0,0023	0,0020	-1,14	0,254	-0,0063	0,0016
E&P dummy	-0,0016	0,0019	-0,85	0,393	-0,0054	0,0021
RMTS dummy	0,0000	(omitted)				
Year 2015 dummy	-0,0635***	0,0130	-4,85	0,000	-0,0892	-0,0378
Year 2016 dummy	0,0099***	0,0032	3,1	0,002	0,0037	0,0162
Year 2017 dummy	0,0024	0,0020	1,17	0,244	-0,0016	0,0064
Year 2018 dummy	0,0000	(omitted)				
Year 2019 dummy	0,0026	0,0026	1,01	0,311	-0,0024	0,0076
Country dummy	-0,0035**	0,0016	-2,18	0,029	-0,0066	-0,0003
Upgrade t	0,0061	0,0037	1,63	0,104	-0,0012	0,0134
Upgrade t+1	-0,0023	0,0033	-0,7	0,483	-0,0088	0,0042
Upgrade t+2	-0,0010	0,0043	-0,24	0,81	-0,0094	0,0073
Upgrade t+3	-0,0062	0,0044	-1,41	0,158	-0,0147	0,0024
Downgrade t	0,0008	0,0047	0,17	0,861	-0,0084	0,0102
Downgrade t+1	0,0049	0,0040	1,23	0,217	-0,0029	0,0127
Downgrade t+2	0,0036	0,0044	0,81	0,42	-0,0051	0,0123
Downgrade t+3	-0,0019	0,0046	-0,4	0,689	-0,0109	0,0072
EBITDA margin	0,0004	0,0037	0,12	0,905	-0,0067	0,0077
Debtratio	-0,0032	0,0068	-0,46	0,644	-0,0165	0,0102
Revenue	0,00000	0,00000	-0,43	0,665	-0,00002	0,00001
Oilprice change	0,0171	0,0178	0,96	0,336	-0,0177	0,0521
Constant	-0,0013	0,0079	-0,17	0,867	-0,0169	0,0142

Appendix 4 - Summary of results from the complete pooled model 2B in dataset (1) with momentum as explanatory variable.

Appendix 5

Fixed effects stock volatility		
Number of obs =		1176
Number of groups =		26
Obs per group:	min =	36
	avg =	45,2
	max =	48
F(16, 25) =		43,39
Prob > F =		0,000
R-sq:		
	within =	0,2822
	between =	0,4528
	overall =	0,0198
corr(u_i, Xb) =		-0,756

Log(volatility)	Coefficient	Robust SE	t	p>|t|	(95% Conf. Intevall)	
Interest rates	-0,1592***	0,0302	-5,27	0,000	-0,2216	-0,0969
Integrate dummy	0,0000	(omitted)				
E&P dummy	0,0000	(omitted)				
RMTS dummy	0,0000	(omitted)				
Country dummy	0,0000	(omitted)				
Upgrade t	-0,1137**	0,0532	-2,14	0,042	-2232	-0,0042
Upgrade t+1	-0,0932	0,0861	-1,08	0,29	-0,2705	0,0841
Upgrade t+2	0,0658	0,0609	1,08	0,29	-0,0597	0,1913
Upgrade t+3	-0,0457	0,0729	-0,63	0,537	-0,1959	0,1045
Downgrade t	-0,0407	0,0799	-0,51	0,615	-0,2054	0,1239
Downgrade t+1	-0,0419	0,0592	-0,71	0,485	-0,1639	0,0799
Downgrade t+2	-0,1074***	0,0388	-2,77	0,01	-0,1872	-0,0275
Downgrade t+3	0,0305	0,0481	0,63	0,532	-0,0685	0,1295
EBITDA margin	-0,1274**	0,0869	-1,47	0,015	-0,3063	0,0516
Debt ratio	0,6118	0,4235	1,44	0,161	-0,2604	1,4839
Log(return)	0,14790	0,118	1,25	0,222	-0,0953	0,3909
Year 2015 dummy	0,0000	(omitted)				
Year 2016 dummy	0,1981**	0,0929	2,13	0,043	0,0066	0,3896
Year 2017 dummy	-0,1577**	0,059	-2,67	0,013	-0,2793	-0,0362
Year 2018 dummy	0,2021***	0,0442	4,57	0,000	0,1109	0,2931
Year 2019 dummy	0,0000	(omitted)				
Oil price volatility	1,0385***	0,1483	7,00	0,000	0,7331	1,3439
Constant	-3,0986***	0,4789	-6,47	0,000	-4,0849	-2,1123
Sigma_u	0,5282					

sigma_e	0,3043					
rho	0,7507					

Appendix 5 - Summary of results from the complete fixed-effects model 1B in dataset (2) with momentum as explanatory variable.

Appendix 6

Random effects stock volatility		
Number of obs =		1176
Number of groups =		26
Obs per group:	min =	36
	avg =	45,2
	max =	48
Wald chi2(19) =		1608,51
Prob > chi2 =		0,0000
R-sq:		
	within =	0,2779
	between =	0,6052
	overall =	0,4225
corr(u_i, X) =		0(assumed)

Log(volatility)	Coefficient	Robust SE	z	p>|z|	(95% Conf. Intevall)	
Interest rates	-0,1584***	0,0303	-5,24	0,000	-0,2177	-0,0991
Integrate dummy	-0,2703**	0,1363	-1,98	0,047	-0,5374	-0,0032
E&P dummy	0,2594	0,1850	1,40	0,161	-0,1033	0,6220
RMTS dummy	0,0000	(omitted)				
Country dummy	-0,1350	0,1094	-1,23	0,217	-0,3494	0,0794
Upgrade t	-0,1197**	0,0509	-2,35	0,019	-0,2195	-0,0199
Upgrade t+1	-0,0994	0,0825	-1,2	0,228	-0,2611	0,0623
Upgrade t+2	0,0514	0,0643	0,8	0,424	-0,0747	0,1774
Upgrade t+3	-0,0637	0,0649	-0,98	0,327	-0,1909	0,0636
Downgrade t	-0,0427	0,0823	-0,52	0,604	-0,2040	0,1186
Downgrade t+1	-0,0460	0,0624	-0,74	0,46	-0,1683	0,0762
Downgrade t+2	-0,1122***	0,0398	-2,82	0,005	-0,1902	-0,0342
Downgrade t+3	0,0294	0,0474	0,62	0,535	-0,0636	0,1224
EBITDA margin	-0,0797	0,0849	-0,94	0,348	-0,2462	0,0868
Debt ratio	0,6540*	0,3502	1,87	0,062	-0,0324	1,3405
Log(revenue)	-0,02580	0,0506	-0,51	0,61	-0,125	0,0734
Year 2015 dummy	0,0000	(omitted)				
Year 2016 dummy	0,1213*	0,0729	1,66	0,096	-0,0215	0,2642
Year 2017 dummy	-0,1985***	0,0491	-4,04	0,000	-0,2948	-0,1022

Year 2018 dummy	0,2020***	0,0451	4,48	0,000	0,1136	0,2904
Year 2019 dummy	0,0000	(omitted)				
Oil price volatility	1,0378***	0,1486	6,98	0,000	0,7465	1,3291
Constant	-2,4244***	0,4163	-5,82	0,000	-3,2403	-1,6085
Sigma_u	0,1959					
sigma_e	0,3044					
rho	0,2931					

Appendix 6 - Summary of results from the complete random-effects model 2B in dataset (2) with momentum as explanatory variable.