The Effect of ESG Performance on Share Price Volatility

Robin Jakobsson, Leo Lundberg

Department of Business Administration
International Business Program
Degree Project, 30 Credits, Spring 2018
Supervisor: Jörgen Hellström
Abstract

Environmental, Social, and Governance (ESG) investing is growing rapidly. Previous research in the area, has mostly been centered around ESG/CSR and its link to corporate financial performance, cost of capital and idiosyncratic risk. Furthermore, relevant previous research is presented that in part challenges the traditional market models and suggests that total risk is a relevant risk factor, instead of only the systematic risk, as proposed by normative theory. In this study, we develop two separate panel regression models, with separate dependent variables. Realized volatility and a GARCH (1,1) estimate of volatility. This is done in order to gain insight into if there is, as propositioned, a negative relation between high ESG/CSR performance and volatility of the shares, i.e. the total risk of the shares. The study uses ESG and financial data from Thomson Reuters Eikon database. The sample size is 481 firms from the S&P 500 Index, for the years 2009-2016. The results of this study indicate that there is a statistically significant negative relationship between high ESG/CSR performance and share price volatility. This result adds to the discussion that challenges existing theory.

Key words: ESG, CSR, corporate social responsibility, total risk, financial risk, idiosyncratic risk, volatility.
Acknowledgements

This thesis was written as the concluding part of the International Business Programme at Umeå School of Business, Economics and Statistics. We would like to thank our families for their support and input. We would like to thank our supervisor, Jörgen Hellström for his guidance and expertise. We would additionally like to thank our patient friends who have given helpful advice and inspiration.

Umeå, 2018-05-15

Robin Jakobsson
Leo Lundberg

robinjakobsson11@gmail.com lelu0032.umu@gmail.com
# Table of Contents

1. Introduction .................................................................................................................. 1  
   1.1 Problem Background .............................................................................................. 1  
   1.2 Research Questions .............................................................................................. 3  
   1.3 Thesis Purpose and Contribution ........................................................................ 3  
   1.4 Delimitations ....................................................................................................... 4  

2. Scientific Method ......................................................................................................... 6  
   2.1 Choice of Subject .................................................................................................. 6  
   2.2 Preconceptions ..................................................................................................... 6  
   2.3 Epistemology ....................................................................................................... 7  
   2.4 Ontological Assumption ...................................................................................... 8  
   2.5 Methodological Assumption ............................................................................... 8  
   2.6 Quantitative and Qualitative Approach ................................................................ 8  
   2.7 Summary of Scientific Method ........................................................................... 10  
   2.8 Ethics in Business Research .............................................................................. 10  

3. Literature Review ........................................................................................................ 11  
   3.1 Empirical Studies on the CSR-CFP Relationship ............................................... 11  
   3.2 Empirical Studies on the CSR-Cost of Capital Relationship ............................. 13  
   3.3 Empirical Studies on the CSR-Risk Relationship .............................................. 14  
   3.4 Summary of Previous Research .......................................................................... 17  
      3.4.1 Studies on CSR-CFP Relation ..................................................................... 17  
      3.4.2 Studies on CSR-Cost of Capital Relation .................................................. 17  
      3.4.3 Studies on CSR-Risk Relation .................................................................... 18  
   3.5 Critique on Previous Research ............................................................................ 18  
   3.6 The Literature Review Process ............................................................................ 20  

4. Theoretical Framework ............................................................................................... 21  
   4.1 Modern Portfolio Theory ...................................................................................... 21  
   4.2 Capital Asset Pricing Model ................................................................................ 22  
   4.3 Efficient Market Hypothesis and Fundamental Analysis .................................. 24  
   4.4 The Portfolio Characteristics of Individual Investors ........................................ 26  
   4.5 Volatility .............................................................................................................. 26  
   4.6 Life Cycle Sustainability Assessment Methodology ......................................... 28  
   4.7 Summary of Theoretical Framework .................................................................. 30  

5. Method ......................................................................................................................... 32  
   5.1 Population and Selection of Sample .................................................................... 32
Table of Tables

Table 1. *Overview of Scientific Method* ................................................................. 10
Table 2. *Summary of Previous CSR-CFP Research* ................................................. 18
Table 3. *Summary of Previous CSR-Cost of Capital Research* ............................ 18
Table 4. *Summary of Previous CSR-Risk Research* ............................................... 19
Table 5. *Descriptive Statistics of Included Variables* .............................................. 43
Table 6. *Example of High Scoring ESG Companies* ............................................. 45
Table 7. *Example of Average ESG Scoring Companies* ........................................ 45
Table 8. *Average ESG Score Development Over Time* ......................................... 46
Table 9. *Correlation Matrix* ..................................................................................... 46
Table 10. *Regression Results A* ............................................................................ 47
Table 11. *Variance Inflation Factor Test* ................................................................. 49
Table 12. *Regression Results B* ............................................................................. 50

Table of Figures

Figure 1. *The Efficient Frontier* .............................................................................. 22
Figure 2. *Overview of LCA and LCC* ................................................................... 30
Figure 3. *ESG Score Overview* ............................................................................ 34
Figure 4. *ESG Score Distribution* ......................................................................... 44
1. Introduction

The aim of this chapter is to introduce the reader to the problem background and research gap that lies as a foundation for this study. This problem background will then lead the reader to the research questions stated in the study and the thesis purpose and contribution. The chapter is ended by delimitations.

1.1 Problem Background

According to the World Wildlife Fund (WWF, 2018) climate change has an effect on many aspects of life on earth, if we are to sustain a growing population, changes are needed. Climate change heavily burdens people and businesses in front line communities, those living nearest to shorelines, floodplains and polluted city-centers. Furthermore, WWF claims that climate change destabilizes food production, displaces people in vulnerable countries, forcing them to relocate. The rift between rich and poor is increasing. There are massive differences in both fairness of labour law and actual working conditions around the world. Executive pay has increased in a much more rapid pace than the pay for the average worker. Corruption is still a major problem for business in parts of the world. (WWF, 2018)

WWF (2016, p. 112) does recognize that solving these issues will require fundamental changes, for example, the way we today measure economic success on a country level is often by GDP. A high and increasing GDP is the goal of governments. The problem with this approach is that it only measures the monetary value created. WWF (2016, p. 112-114) suggests that we should replace this old approach by measuring economic, ecological and social performance, all actors are urged to take a more long term approach and avoid short-termism. Furthermore, WWF (2016, p. 112) notes that redirecting financial flows in a way that enables sustainable ecosystem management is of essence. WWF (2016, p. 112) continues by recognizing that there is a cost that comes with the resource scarcity that is caused by ignoring sustainable business. WWF (2016, p. 112) claim that businesses should also want to avoid the cost that may come with increased regulation because of deteriorating climate or due to reputational disasters.

At the same time, finance is undergoing a small revolution. An example of this is that Environmental Social and Governance (ESG) investing has risen immensely in importance during recent years, to a large extent as a response to the fact that people are getting more and more aware of the issues mentioned. The Global Sustainable Investment Alliance (GSIA) identify a number of different strategies that investors use when investing sustainably. Investing based on ESG Scores or criteria is according to GSIA the second most used strategy, being trumped only by negative or exclusionary screening. Furthermore, ESG integration in some form, as an investment strategy is stated to encompass 10.37 trillion USD (Global Sustainable Investment Review, 2016, p. 3). Only between the years 2014 and 2016, the total amount of assets invested in socially responsible projects has increased by roughly 25%, from $18.28 to $22.89 trillion (Global Sustainable Investment Review, 2016, p. 7). The GSIA define ESG integration as “the systematic and explicit inclusion by investment managers of environmental, social and
governance factors into financial analysis” (Global Sustainable Investment Review, 2016, p. 6). Previous research in the field of sustainable investment is heavily weighted toward exploring relationships between Corporate Social Responsibility (CSR) performance and Corporate Financial Performance (CFP), such as studies by Wang & Sarkis (2017), Rodriguez-Fernandez (2016), Filbeck et.al. (2009), Brammer et.al. (2006), Guenster et.al. (2011), Schnietz & Epstein (2005), Hillman & Keim (2001), Turban & Greening (1997) and McGuire et.al. (1988). Another bulk of research has gone into studying the relationship between CSR and cost of capital, such as in Suto & Takehara (2017), Cajias et.al. (2014), Chava (2014), El Ghoul et.al. (2011) and Sharfman & Fernando (2008). As well as research that has gone into studying the relationship between CSR and idiosyncratic risk, such as Nguyen et.al. (2015), Mishra & Modi (2013), Humphrey et.al. (2012), Jo & Na (2012), Lee & Faff (2009), Luo & Bhattacharya (2009) and Spicer (1978). ESG and CSR are two closely related disciplines. The European Commission (EC) highlights that CSR should be company-led. Measures that companies can take are among others, “following the law, integrating social, environmental, ethical, consumer, and human rights concerns into their business strategy and operations.”(European Commission, 2018). According to one of the leaders in sustainable asset management, Robeco SAM (2018), the environmental factor of ESG include aspects such as for example, the impact a company has on climate change, the waste management process and the use of energy within companies. The social factor includes, but is not limited to, aspects such as how well a company respects human rights, avoid child labour and adhere to workplace safety and health. Furthermore, Robeco SAM (2018) state that the social factor also includes how well integrated companies are in their respective local communities and how the local communities are being respected. The governance factors include, but are not limited to, aspects such as defining rights, responsibilities and expectations between different stakeholders in the governance of firms. A well functioning corporate governance supports a company in its long-term strategy. The quote by the European Commission highlights the CSR aspects, while Robeco SAM point out the ESG aspects. As can be seen, the issues that both ESG and CSR handle are broadly the same. This highlights the fact that the difference between ESG and CSR is mostly the perspective taken. ESG aspects are the concern of investors, while CSR is something that is conducted on the business level. So, while having a CSR policy, doesn’t automatically translate into high ESG performance, it is a starting point. However, a good CSR policy should by all logic translate into good ESG performance, since the same aspects are evaluated. In this thesis ESG and CSR are used interchangeably, since the main difference between the two is the perspective one takes in the analysis.

Some of the early empirical research on financial risk is concerned with the risk-return tradeoff, such as the Modern Portfolio Theory (MPT), originally coined by Harry Markowitz in 1952. Subsequent work by British statistician Maurice Kendall introduced the Efficient Markets hypothesis in 1953. In the 1960’s the Capital Asset Pricing Model (CAPM) was introduced as a simplification and development on the work of Markowitz’s MPT. The CAPM was introduced after various contributions by Jack Treynor, John Lintner, William Sharpe and Jan Mossin. These models are considered to be traditional models in finance, what they all have in common is that they are concerned with systematic risk only. Idiosyncratic risk does not matter according to these theories. The two components of volatility, or total risk are systematic and idiosyncratic risk. CSR aspects are managed on company level, and as such should, if anything, affect the idiosyncratic risk. Early research in the area seems to be scarce, but recent research by various authors, such as Campbell et.al.(2001), Wei & Zhang (2006), Kumar & Lee
(2006), Ang et.al. (2006), Malkiel & Xu (2003) and Fu (2009), suggest that idiosyncratic risk has increased as compared to systematic risk, at least periodically. Furthermore the studies suggest that idiosyncratic risk should matter, due to their various findings. Another bulk of research investigate the portfolio characteristics of individual investors, such as Barber & Odean (2000), French & Poterba (1991), Goetzmann & Kumar (2008), Grinblatt & Keloharju (2001), Polkovnichenko (2008) and Kilka & Weber (2000). The findings point to characteristics such as investors portfolios not being diversified enough to be completely without idiosyncratic risk. In order to investigate the relationships mentioned, some sort of measurement model needs to be used for both the risk aspect and the CSR aspect. The simplest form for measuring total risk is realized total risk, or realized volatility. More advanced models include approaches such as for example General Autoregressive Conditional Heteroskedasticity (GARCH) models. In this study, versions of these two models will be used to measure volatility. In previous research many different databases have been used for measuring CSR. In this study, the Thomson Reuters ESG Score will be used. This is a comprehensive measure for how well companies conduct CSR.

Previous research into the field of CSR-risk relationship has largely discovered a negative relation between the variables. This is the case in studies by Harjoto et.al. (2017), Mishra & Modi (2013), Jo & Na (2012), Lee & Faff (2009), Luo & Bhattacharya (2009) and Spicer (1978). While there also are studies pointing to no relation, or a positive relation, such as Humphrey et.al. (2012) and Nguyen et.al. (2015) respectively, the previous research points to a negative relation, on balance. Given the general large and growing interest in ESG investing, and the fact that previous research in ESG and CSR focuses mainly on CSR-CFP, CSR-cost of capital and CSR-idiosyncratic risk. There seems to be a research gap in the field of CSR/ESG performance and its effect on total risk, as only one study has been found in this field (Harjoto et.al., 2017). The previous research presented here leads us as authors to be inclined to expect a negative relation between ESG/CSR performance and volatility in this study as well. Subsequently, the following section will present the research questions formulated to gain insight into this research gap.

1.2 Research Questions

Research Question 1 (RQ1): Does a high ESG-score lead to lower share price volatility, as measured by daily GARCH (1,1) estimates of volatility, ceteris paribus?

Research Question 2 (RQ2): Does a high ESG-score lead to lower share price volatility, as measured by simple realized daily volatility, ceteris paribus?

1.3 Thesis Purpose and Contribution

The purpose of this thesis is to study the effect of ESG-performance on total share price volatility. The ESG score variable will be retrieved from Thomson Reuters’ EIKON. This will be done by using two different methods of estimating historical volatility. A simple form of daily realized volatility measure as well as a GARCH (1,1) model of volatility estimation will be used to answer the research questions. ESG-performance as the authors use the term, includes the management of environmental, social and governance factors. In previous research, it has been commonplace to use only one method of volatility estimation, such as in Filbeck et.al. (2009), Brammer et.al. (2006), Guenster et.al. (2011),

Previous studies in the field of ESG have been heavily weighted toward investigating the effect of CSR on firm profitability such as in, Filbeck et al. (2009), Brammer et al. (2006), Guenster et al. (2011), Schnietz & Epstein (2005), Hillman & Keim (2001), Turban & Greening (1997), McGuire et al. (1988), or on firm idiosyncratic risk such as in, Mishra & Modi (2013), Humphrey et al. (2012), Jo & Na (2012), Lee & Faff (2009), Luo & Bhattacharya (2009) and finally the effect of CSR on firm cost of capital (Suto & Takehara, 2017; Cajias et al., 2014; Chava, 2014; El Ghoul et al., 2011; Sharfman & Fernando, 2008) The measures used in this study are for total share price volatility. This issue is much less studied, and the authors have only been able to find one such study (Harjoto et al., 2017).

Given the portfolio characteristics of individual investors, additional insight into how good ESG performance affects total risk is of interest. The result of this study will lead to the authors being able to make recommendations to individual investors, concerning how ESG performance affects total risk and how this can be used in portfolio formation, with certain assumptions.

Further, this study is also valuable from a firm-perspective. The research will highlight the benefits of including CSR activity in the corporate agenda. Insight into how the shares of a firm behave when an effective CSR policy is in place is of value in the debate of whether there is merit in focusing on CSR activity, from a company perspective. The rift between short-term profit seeking and long-term CSR activity is discussed, in order to make companies aware that there are issues, which when managed correctly may let companies gain the benefits that seem to arise from good CSR/ESG performance.

1.4 Delimitations

The literature review has been restricted to include previous studies concerning the CSR-risk, the CSR-corporate financial performance (CFP) relation as well as the relation between CSR and cost of capital. The reason for this is that the timeframe of this study is limited, and while there is more research within the area of CSR, the most relevant research for this study has been chosen to be these three aforementioned areas.

The chosen dataset on ESG data from Thomson Reuters is provided only once every year, while the returns are calculated on daily share closing prices. Comparing a variable measured on a daily level with other variables on yearly level is usually suboptimal. Mitigating actions could have been taken to address this issue, in the form of recalculation of the returns to perhaps monthly returns, or even quarterly or even yearly, to have the data on the same level of frequency. The reason for the fact that this wasn’t done is that the authors believe that ESG Scores are expected to be very stable over the period in which they are reported. This is reasonable because of the nature of ESG investing and ESG work, it is very much a long-term issue, Peloza notes this in a 2006 article as well. Both results and investment results happen with persistent, dedicated work. Companies that put in this effort are also, as a result expected to have relatively stable ESG performance over time. Another aspect that speaks for this view is that the reputation of a company is likely
to be forged over a longer period of time. For investors to notice the ESG work being done by a company, there needs to be a long-term mindset present, so it is reasonable to believe that it is in the best interest of each company to strive for a stable, high ESG performance. The collected data backs these claims up. This can be seen in Table 6, Table 7 and Table 8. The ESG Scores generally remain stable, anchored on the underlying CSR performance of the companies. Furthermore, using yearly measures for volatility would lead to a loss of accuracy in the measurement of volatility. The higher frequency measure includes more information about estimated and realized volatility. These arguments lead the authors to believe that it is, in this case, reasonable to compare daily variance with a yearly ESG score variable.
2. Scientific Method

This part will address the choice of subject, as well as the preconceptions held by the authors about this subject. Furthermore, there are several assumptions made in regards to scientific methodology which will be clarified. Last, business research ethics will be handled and a brief summary of the chapter is presented.

2.1 Choice of Subject

The authors of this study have an interest in finance, and have accordingly chosen this specialization for their master's program. The subject of sustainable investments has been of interest to the authors of this study during the time at Umeå University. Different research topics involving sustainable investments were discussed. One such topic was the much more researched area of corporate social responsibility and financial performance. The authors, Jakobsson and Lundberg, originally wanted to conduct this study with data from companies residing in Sweden, or the Nordic countries (Sweden, Norway, Finland, Denmark and Iceland). This was considered to be of interest mainly because these are the markets where the authors live, but also because studies of this character were, to the authors knowledge, non-existent in this region. The Thomson Reuters EIKON database, responsible for providing the ESG scoring system, which the data for the study is taken from, was upon investigation deemed to not be able to provide good enough ESG data on Nordic companies. There were relatively few Nordic companies evaluated by Thomson Reuters’ ESG score, with far from complete annual records for all of them. This made the authors of this study look at the availability of ESG scores in other markets within the same database. The S&P 500 companies within the Thomson Reuters’ ESG database were much more extensive, as all 504 companies from that U.S. market index were scored, at least some years. Upon first glance, it was estimated that only a few of these would not be able to provide ESG scores for the years 2009 to 2016. This was considered a satisfactory dataset to base the study on.

2.2 Preconceptions

It's inevitable for authors of a study to not have preconceived ideas about the topic being handled. Authors can be more or less aware of these preconceptions influencing the study as it progresses. According to Bryman & Bell (2011, p. 40-41), these inherent preconceptions from the authors will naturally not be possible to get rid of, which is why it is important to highlight any possible biases which can influence how a study is presented. When conducting a scientific study, the authors can bring in certain biases which risk to affect the outcome of the study. These biases tend to originate from beliefs and previous experiences. In the context of a quantitative study, and more precisely validity, this can express itself in the way of authors choosing to present their findings and data in a limited manner (Bryman & Bell, 2011, p. 40-41).

The authors, Jakobsson and Lundberg, are aware of environmental, social and governance issues for companies in general, and has an interest in its causes and effects. How a company's shares behave, as a result of how a company handles these issues, is a topic
which has been present in the minds the authors for some time. In the courses of the Civilekonom program at Umeå University, it has been touched upon the growing importance of ESG/CSR, especially as environmental issues become more relevant each day.

The authors of this paper have very limited experience in working within the field of finance and are yet to get their first jobs in the area. This brings with it both strengths and weaknesses. The inexperience may lead the authors to making conclusions that are not viable in real-world applications. The ability to reflect over the issue at hand is limited to knowledge acquired from textbooks, financial press, accompanied by the overall business sense of the authors, this might lead to a certain lack of wholeness in the analysis as the inexperience in the field is certain to be reflected somehow. The inexperience may however also allow the authors to analyze this issue without any preconceived notion about how things are. This may give the authors the chance to view the issue from a more unbiased perspective.

2.3 Epistemology

According to Saunders et al. (2012, p. 132), epistemology is concerned with the issue of what should be considered valid knowledge within an area of study. A central question is whether social science should adhere to the same methods as those used when performing a natural science study (Bryman & Bell, 2015, p. 26). It is important to clarify which paradigm the authors of a study have chosen, in order to avoid misunderstandings at the most basic level of research methodology. There are two main paradigms, positivism and interpretivism. Positivism takes the stance that only phenomena that are observable and measurable can be considered valid knowledge. An objective and independent way of handling what is being studied is positivism. Interpretivism is of the view that the distance between researcher and what is being studied should be minimized, as to participate in close inquiry (Collis & Hussey, 2014, p. 43-46).

According to Collis & Hussey (2012, p. 44), positivism is mainly associated with natural science. According to Saunders et al. (2012, p. 134), only phenomena that can be observed will result in the production of credible data, if adhering to a positivist paradigm. Collis & Hussey (2014, p. 50-53) points out that for a long time, this was the only paradigm, as science was up to a time in history mainly concerned with measuring inanimate objects in the physical world. From observation and experiments, inductive logic was used to discover theories which could be used for prediction (Collis & Hussey, 2014, p.50-53). Interpretivism is mainly associated with social science, according to Collis & Hussey (2014, p. 44-45). It tends to have a natural location, use small samples, allow findings to be generalized from one setting to another similar setting, be concerned with generating theories, produce rich, subjective, qualitative data and produce findings with low reliability and high validity (Collis & Hussey, 2014, p. 50). Saunders et al. (2012, p. 137) emphasizes an important thing to understand about interpretivism, which is that the researcher has to adopt an empathetic stance. This can roughly be translated to entering the social world of the research subjects, and understanding matters from their point of view. This study takes a positivist approach, since the characteristics of natural science are present through most parts of the process. The use of a predetermined approach in testing a set hypothesis followed by a statistical analysis of numerical data is considered positivist.
2.4 Ontological Assumption

Ontology is concerned with the nature of reality. According to Collis & Hussey (2014, p. 47), the objectivist view is that there is a reality independent of human interference. The subjectivist view is that reality is created by people, which leads to facts being a construction of human experience. Bryman & Bell (2015, p. 32-34) explains that objectivism takes the ontological position that social phenomena are external facts beyond our reach and influence; an organization, such as a large company, is viewed as having the characteristics of an object. It is possible to view it from the outside, with its processes and structures. Subjectivism is of the view that social interactions, such as hierarchy, are worked out naturally, which causes any imposed organizational structure to clash with the 'real' structure, that individuals have no role in fashioning. 'The organization' is a socially constructed product, a label to make sense of their experience.

The ontological assumption of this study is that of objectivism. This is because the authors Jakobsson and Lundbër have no intention of employing a subjective viewpoint at any point in this thesis. The first part of the quantitative data for the study is undoubtedly without subjective tendencies, which is the daily returns of companies used. This is simply based on the daily closing price of each company's listed shares. The second major part of the study is the quantitative ESG score used to rate to which degree companies engage in sustainable investment. According to the company producing said score, Thomson Reuters, their metrics for scoring companies are thorough and trustworthy; this is however beside the point, as the ontological assumption concerns the authors of this study (Thomson Reuters, 2018). The control variables are also handled objectively.

2.5 Methodological Assumption

Deductive theory is, according to Bryman & Bell (2015, p. 23-26), the most common way to relate research and theory. Saunders et al. (2012, p. 145) describes the deductive approach as a large part of what we think scientific research is. The researcher creates a hypothesis, based upon the knowledge of a subject and its theories, which must then be subjected to empirical scrutiny. In the hypothesis will be concepts that need to be translated into researchable entities. Moreover, Bryman & Bell (2015, p. 23-26) also points out that if the finding of a deductive study stands in contrast to what the original theory suggests, the theory needs to be revisited and possibly modified; this part is called inductive research. Deductive research does not necessarily exclude inductive research. This is a rough generalization of the relationship between theory and research when it comes to deductive and inductive strategies. However, it is perhaps better to think of deductive and inductive strategies as tendencies rather than absolute distinctions. A deductive approach is used in this thesis.

2.6 Quantitative and Qualitative Approach

According to Bryman & Bell (2015, p. 37-39), a quantitative research strategy stands as a counterpart to qualitative research. Quantitative research is mainly associated with positivism, and entails working with data which can be quantified, meaning it can effectively be translated into numerical values. It is associated with a deductive approach to the relationship between theory and research, which emphasizes testing theories. A
quantitative research strategy takes the view of social reality as objective and external, meaning values derived will not be subjected to interpretation based on which social context it is placed in. Saunders et al. (2012, p. 472) point out that quantitative data, in its raw form, conveys very little meaning to most people. The data has to be processed and analyzed in order to be able to overview it and derive conclusions from it.

Qualitative research will in contrast be difficult, if not impossible, to translate into numerical values. Even if coding the qualitative data into numericals would be possible, it would per definition stop being qualitative data at that point. Because of this it is fair to say that qualitative research often uses words rather than numbers to describe what is being studied. It takes the view of social reality as a constantly changing emerging property of an individual's creation. The qualitative approach has rejected the practices and norms of the natural scientific model, and that of positivism in particular, in favour of putting the individual's interpretation as central. Moreover, it is heavily leaning towards the inductive approach to the relationship between theory and research, where, according to Bryman & Bell (2015, p. 37-39) the emphasis is on generating theories.

Collis & Hussey (2014, p. 54-55) explain that describing feelings of happiness or sadness could either be done through an interview with the person, making it a qualitative study, or possibly through a brain scan at a hospital, which would yield quantitative results. Naturally the brain scan would produce highly accurate and objective results, with a procedure which could be applied to many people to yield comparable results. It is however unclear if this would give us more useful information about happiness or sadness compared to doing a qualitative study, where people are interviewed about the matter. It depends on what exact problems you are hoping to bring clarity about. According to Collis & Hussey (2014, p. 54-55) there is not necessarily anything prohibiting a study from incorporating both quantitative and qualitative data methods. Studies making use of both quantitative and qualitative data methods are considered to apply pragmatism. This approach can prove useful when a single paradigm is constricting the selection of methods to effectively answer the research questions at hand (Collis & Hussey, 2014, p. 54).

This thesis adheres to a quantitative method of research. Determining if sustainable investments has an effect on share volatility has to be done through a statistical analysis to yield a credible result. Share price volatility has to be measured numerically, with the most common measure being the standard deviation or variance of returns. The amount of sustainable investments a company engages in has to be quantified in order to be used in a statistical analysis.
2.7 Summary of Scientific Method

Table 1. Overview of Scientific Method.

<table>
<thead>
<tr>
<th>Area of research design:</th>
<th>Chosen approach:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epistemological Assumption</td>
<td>Positivist</td>
</tr>
<tr>
<td>Ontological Assumption</td>
<td>Objectivist</td>
</tr>
<tr>
<td>Methodological Assumption</td>
<td>Deductive</td>
</tr>
<tr>
<td>Data Collection</td>
<td>Quantitative</td>
</tr>
</tbody>
</table>

2.8 Ethics in Business Research

Bryman & Bell (2007), cited by Collis & Hussey (2014, p. 31-34) identify principles of ethics that should be adhered to. In this section some of the principles most applicable to this study will be described, the issue and how the authors of this paper have acknowledged them will be discussed later on in the paper.

The seven principles are:

- **Harm to participants** is “the need to avoid potential harm through the research process and the need to ensure physical and psychological wellbeing of research participants, the researcher and others”.
- **Informed consent** is “the need to ensure the fully informed consent of research participants”
- **Privacy** is “the need to protect privacy of research subjects or avoid invasions of privacy”
- **Deception** is “the potential for deception during the research process, either through lies or behaviour that is misleading”
- **Affiliation**, “the need to declare any professional or personal affiliations that may have influenced the research”
- **Honesty and transparency** is “the need for openness and honesty in communicating information about the research to all interested parties, including the need for trust”
- **Reciprocity**, “the research should be of mutual benefit to researcher and participants or some form of collaboration or active participation should be involved”
- **Misrepresentation** “the need to avoid misleading, misunderstanding, misrepresenting or falsely reporting the research findings” (Collis & Hussey, p. 31-34)

In this study, the authors will try to adhere to these principles. The risk of harm to participants, the issue of informed consent and privacy is minimized through using secondary, publicly available data, making these issues inherently small. The authors do not intend to attempt to deceive in any way, and data will be presented as truthfully as possible. In section 2.2 the authors state their relationship to the subject, in order to clarify for honesty and transparency. The authors have no affiliation with Thomson Reuters or any outside agents. The research is intended to lead to increased insight into the subject studied, leading to a reciprocal benefit for society, companies and to the authors and readers of this study. The authors have put in their best efforts in trying to avoid misrepresentation of any kind in the research process.
3. Literature Review

So far, research on the subject of CSR has involved several different topics. The most pressing questions have been if it is at all profitable to engage in CSR measures for a company, the effect of CSR on cost of capital, as well as the relation between CSR and risk. In this chapter, these three areas of previous research will be presented. The main results of previous research will then be summarized. Critique on previous research and the literature review process will be presented.

In order to position our study in accordance to previous research and to provide for a foundation to discussing the implications of the results, some of the findings of previous research will here be outlined. Again, the reason that these are studies investigating "CSR" and not "ESG" specifically, is that ESG research is done mostly by financial institutions to provide for their clients. CSR research, on the other hand is conducted by researchers. Peer-reviewed articles on the subject of sustainability generally refer to the sustainability aspects outlined in section 1.1 as CSR aspects.

The three areas of previous CSR research included in this study concern the areas of CSR and its relation to corporate financial performance, cost of capital and risk. In order to gain insight into the research question and purpose, it is noted that these three major areas overlap and have implications for a broader impact of a change in risk on the company level. That is, a possible effect of ESG performance on risk, would also render interesting results, when paired with this bulk of previous research. The risk-return trade-off in finance, as well as the fact that risk is a determinant of cost of capital, intertwine these areas of research.

3.1 Empirical Studies on the CSR-CFP Relationship

Wang & Sarkis (2017) research the relationship between CSR and corporate financial performance (CFP). The CSR data is retrieved from the Bloomberg ESG database and financial data is retrieved from Compustat. Compustat is a very extensive database by Standard & Poor, which provides financial data on businesses. The sample of the study are US firms from the top 500 Green companies for the years 2009 to 2013. It is found that CSR outcomes have a mediating role in the relationship between CSR and CFP. Furthermore, Wang and Sarkis (2017) indicate that the implementation of the CSR aspects into the business is of essence in order for the CSR practices to influence corporate financial performance positively. In order for CSR to work, the authors argue that a rigorous and real effort to integrate CSR into all areas of business eventually leads to better financial performance. It is however not enough for companies to merely “window-dress” their CSR work, on the contrary, the effort has to be genuine (Wang and Sarkis, 2017, p. 1607, 1615).

Rodriguez-Fernandez (2016) investigates the bi-directional relationship between financial performance and CSR using data for 121 companies on the Madrid Stock Exchange, for the year 2009. For the CSR data component, several different measures was employed to produce a combined measure. The data used for the CSR component
was the GRI index (sustainability reporting, independent organization), the Dow Jones Sustainability Index and the Global Compact Network. A regression analysis was carried out to see if 1) companies displaying greater CSR behaviour leads to improved financial performance, and if 2) improved financial performance leads to better behaviour in CSR. Both of these assumptions were deemed to be true, meaning that CSR and financial performance both effect each other, creating a virtuous circle.

Filbeck et.al. (2009) study companies that perform well in the best companies for leaders. The survey is done in the Chief Executive annual survey. The study uses data on US firms. The news announcement impact is studied for the firms included and the holding period is the period of subsequent survey releases. It is found that the leaders outperform when raw and risk-adjusted returns are considered, during times of high volatility.

Brammer et.al. (2006) on the other hand find a negative link between CSR performance and returns, by studying UK firms. The lower the performance in social aspects, the higher the return was proved to be. Only one social performance indicator was in use in the study. Industry effects were not able to explain the poor performance of high social performers. The Ethical Investment Research Service (EIRIS) is used to access data on CSR aspects. The database is based on a survey, but EIRIS does their own research as well. The sample used was all companies that comprise the FTSE All-Share Index, this should give a good spread of companies across industries.

Guenster et. al. (2011) look into if “eco-efficiency” has a positive impact on financial performance. The variable is described as how well a company manages the waste it creates by environmentally friendly measures. A positive relation was found between eco-efficiency and financial performance in the study. Return on assets and Tobin’s Q are used as variables for financial performance. The variable eco-efficiency has a survey as its major data collection method. The scores for eco-efficiency used in the study are produced by Innovest Strategic Value Advisors. Portfolios were developed based on the measure and the high ranked portfolio outperformed the lower ranked one during the study period of 1997-2004. The difference could not be explained by market sensitivity, investment style or industry specific factors. Regression analysis was the main statistical method used and firms from the US stock markets are included.

Schnietz & Epstein (2005) study the value of CSR in times of crises, the study focuses on US companies, more specifically fortune 500 companies. The 1999 crisis is the focus in the study, which is formed as an event study. The empirical model used is a regression analysis and the Domini Social Index mutual fund is used as a proxy for sustainability. There were a total of 416 companies in the sample. The study found that CSR performance lead to better financial performance in crisis, and this while checking for industry and trade effects.

Hillman & Keim (2001) use data on the S&P 500 in order to study if their variable “stakeholder management” and “social issue participation” lead to added shareholder value. This is done using a regression analysis of the years 1994-1996. The variable stakeholder management performance is retrieved from the Kinder, Lydenburg, Domini database. It includes 9 measures on stakeholder management, and can be described as a less extensive data source than the Thomson Reuters ESG Score. The study finds that both “stakeholder management” and “social issue participation” are positively related to shareholder value. The variable “market value added” (proxy for added shareholder
value) is retrieved from Stern Stewart Performance 1000 database, which tracks the Fortune 1000.

Turban & Greening notes in a 1997 publication, which examines corporate social performance and organizational attractiveness for potential employees, that firms with a strong CSR commitment often have an increased ability to attract and retain employees. This leads to reduced costs due to its impact on recruitment, training and turnover. In the study, the authors expected larger and more profitable firms to have a better name recognition, which made them implement the control variables firm size and profitability. The result indicated that firm size and profitability were not the factors making these companies more attractive, rather it was indicated that the strong CSR commitment lead to this conclusion (Turban & Greening, 1997). The sample for this study was drawn from the KLD database. While this study does not examine financial performance directly, it does indicate the possibility for strong CSR commitment to reduce costs, which is positive for the bottom line of a company.

McGuire et.al. (1988) use data from the Fortune 500 survey on corporate reputations and was conducted in the USA, in order to try to find a link between corporate social responsibility (CSR) and financial performance. The survey in question had over 8000 respondents, these were executives, outside directors and analysts. The response rate was 50% in the survey. The study finds a positive relationship between CSR and the accounting based measures used. The authors of this study state that the major shortcoming of the study is that there is a problem with measuring CSR, both in the length of the available data and substance.

3.2 Empirical Studies on the CSR-Cost of Capital Relationship

Suto & Takehara publish a 2017 study, focusing on the Japanese market between the years 2007 to 2013. The object of study is to find out if there is a relationship between CSR activity and cost of capital, while checking for the influences of banking relationships and ownership structure. The study examines the corporate social performance (CSP) and cost of capital relation in terms of cost of debt, cost of equity and weighted average cost of capital. The CSR measure is one aggregated measure and is based on an annual survey done by Toyo Keizai Inc. A regression model is used for analysis. The findings point to that institutional ownership influences the CSP - cost of equity relation, reducing cost of equity. Between the years 2008 and 2010 the relation between CSP and bank dependency increases the cost of debt, the effect is diluted between the years 2010–2013 when Japan moved to a more market based economy, according to Suto & Takehara (2017).

Cajias et al. (2014) conducted a study to investigate if continuous investment in CSR had an effect on cost of capital for US companies. The study used a sample of 2356 listed US companies for the years 2003-2010. The KLD database was used for CSR ratings, and financial data was taken from Thomson Reuters Datastream and the Kenneth R. French database. The findings of the study indicate that both capital markets and investors pay higher premiums for firms allocating resources to sustainable activities. An interesting detail found in the study is that companies with high a number of CSR issues, which is to say issues which could be solved by increased CSR activity, had decreased cost of capital in the short run, but suffered a reverse effect in the long run.
Chava (2014) studies the impact of a firm’s environmental profile on the firm’s cost of equity and debt capital. Implied cost of capital is derived from analysts’ earnings estimates. It is found that investors demand significantly higher returns from companies excluded by environmental screens, such as excluding hazardous chemical and large emissions concerns, than from companies without similar concerns. Cost of debt is found to be higher for firms with environmental concerns. KLD data is used in order to get information on the environmental profile of the companies. Financial data is mostly gathered from Compustat and CSRP. US firms form the sample for the study. A total of 5879 bank loans are analyzed between the years 1992-2007 in order to be able to investigate the effect on cost of debt. The sample includes over 500 firms, from US stock markets.

In a study published in 2011, El Ghoul et al. examine the relationship between CSR and cost of equity capital for US firms. Four databases were employed in the study: KLD STATS by KLD for CSR data, Thompson Institutional Brokers Earnings Services for analyst forecast data, Compustat for financial data, and CSRP monthly return files for information on stock returns. The sample is large, consisting of 2809 companies for the years 1992-2007. The number of observations available for these companies increased over time, with a peak in 2004. Four different models were used to calculate the cost of equity capital for the companies, which then had the 10-year US treasury bond yield rate subtracted from results. The findings of El Ghoul et al. (2011) show that firms with better CSR scores have access to cheaper equity financing. Two industry sectors, tobacco and nuclear power, were found to have significantly higher cost of equity.

Sharfman & Fernando find, in a 2008 study that by improving environmental management, the cost of capital can be lowered. Furthermore, the study found that improvements in environmental risk management leads to benefits such as, reduced cost of equity capital, being able to shift from equity to debt financing and a tax reduction by being able to add debt, creating an interest tax shield. The sample is 267 firms from the USA. The data for environmental management is gathered from two separate sources. TRI data, which is on how well companies adhere to the standards set forth by the environmental protection agency. The second source of data is Kinder, Lydenburg and Domini (KLD). The financial data used in the study is gathered from Bloomberg.

3.3 Empirical Studies on the CSR-Risk Relationship

Harjoto et.al. (2017) research the relation between CSR and institutional ownership, and the impact of this on stock return volatility. It is found that the relation between CSR and institutional ownership is a concave function and institutional investors alter their holding in a stock, when the perceived optimal investment in CSR is surpassed. The mediating effect on stock return volatility is also studied and the results show that CSR decreases stock return volatility but at a decreasing rate through its effect on institutional ownership. The results remain robust to different CSR measures and estimation methods. In other words, CSR activity reduces volatility to a certain point, but investment in CSR above this seems to increase volatility again. The CSR data is retrieved from the Kinder, Lydenburg, Domini Database (KLD). The data on institutional ownership is retrieved from the Thomson Reuters 13F database. The years under investigation in the study are 1994 to 2012 and firms from the United States have been used.
Nguyen et al. (2015) conducted a study which investigates the link between CSR and risk. The sample is taken from all the firms of the S&P500, which are rated by the KLD, for the years 1991 to 2003. This results in 3728 firm-year observations. The data is taken from the KLD, which generates an annual evaluation of social performance for a number of different categories. The categories are: community, diversity, employee relations, environment, product, human rights and corporate governance. Financial data is taken from Compustat, and stock prices used in calculating firm value is taken from the Center for Research in Security Prices (CSRP). The findings of the study indicate that there is a positive relationship between CSR and firm risk. When committing to a more favourable treatment of employees, the company incurs a fixed cost which will transfer risk to shareholders.

Mishra & Modi (2013) look into the effect of both positive and negative CSR on firm idiosyncratic risk. CSR data is obtained from the commonly used Kinder, Lydenburg and Domini (KLD) database. Positive CSR is defined as a measure of strengths that firms have. While negative CSR is defined as being able to avoid negative CSR events, or addressing CSR concerns. Both of these aspects are measured in the KLD database. 192 firms from the USA are included in the sample and the years 2000-2009 are included in the study. The authors find that positive CSR decreases firm risk and negative CSR increases firm risk, as hypothesized.

Humphrey et.al. (2012) investigate how corporate social performance (CSP) affect the firm cost of capital and risk. The study finds no difference in risk-adjusted performance of the UK firms with high and low CSP ratings. There is no difference found in idiosyncratic risk either. The study suggests that CSP investments neither create nor destroy value for shareholders. High ranking CSP portfolios seemed to have lower betas, however this finding did not hold up in robustness tests. The study is based on detailed, proprietary ESG ratings data from the Sustainable Asset Management Group (SAM), for the period 2002 to 2010. The UK firm portion of SAM’s data includes 256 companies. The methodology that SAM uses in their ESG rankings closely resembles the methodology of Thomson Reuters, the data can be considered of the highest quality. The study notes that the SAM database is likely to be of higher quality than previously used databases, such as EIRIS or the Kinder, Lydenburg and Domini database. The study is conducted around the time of the latest financial crisis. This is addressed by studying the years 2002 to the beginning of 2008 separately. There is no relation between risk and CSP here either. It is concluded that UK markets seemed efficient in integrating ESG information between the years 2002 and 2010.

Jo & Na (2012) measure the effect of CSR on firm risk in controversial industry sectors. Two mutually exclusive hypotheses are developed, that the risk is reduced, or that CSR is merely window-dressing. The sample is US firms between the years 1991 and 2010. Controversial industries are among others, gambling and tobacco industries. It is found that CSR performance has a negative relation to firm-specific risk, after controlling for various firm characteristics. The window-dressing hypothesis is not supported, but the risk reduction one is. ESG scores provided by MCSI (former Kinder, Lydenburg & Domini database) are used in the study, while remaining data is gathered from COMPSTAT and CRSP.

Lee & Faff (2009) study the effect of corporate social responsibility on idiosyncratic risk. This is done by creating portfolios of leaders and laggars in CSR performance, the leaders
are put in different portfolios than the laggards, then the portfolio with leaders is compared to the laggards. The Dow Jones Sustainability Index (DJSI) is used as a proxy for CSR performance. The DJSI measures five major types of corporate sustainability. These are strategy, financial ability to meet shareholders demands, the third aspect is to foster loyalty by investing in customer relations and a general focus on the long term in operations. Fourth, demands the highest standards in corporate governance and reporting. Fifth, is the aspect of managing human resources and foster satisfaction of the workforce. The primary source for this is a questionnaire. A regression analysis is used to determine the relationships. It is found that leaders in CSR performance have lower idiosyncratic risk than the laggards do. The DJSI is based on the Dow Jones Global Index (DJGI). The DJGI constitutes 5000 companies, while the DJSI contains 2500 companies, the sample is then drawn from the DJSI, approximately 10% of the DJSI companies were included in the study.

Luo & Bhattacharya (2009) find that corporate social performance leads to lower idiosyncratic risk. The data on corporate social performance is retrieved from Fortune Magazines survey, over 10,000 executives, analysts and managers are interviewed on their subjective views of the companies corporate social performance. The rest of the data in the study was obtained from COMPSTAT and The Centre for Research in Security Prices (CRSP). The data in the study covers the years of 2002 and 2003 and is done on US firms. There were eight control variables in the study. These were, profitability, profits volatility, leverage, market-to-book ratio, market capitalization, dividend pay, firm age and firm diversification. Idiosyncratic risk is measured from a basis of daily returns data.

In a study from 1978, Spicer (1978) investigates the relationship between pollution control and profitability, size, total and systematic risk, as well as price/earnings ratio. The data was taken 18 companies in the pulp and paper industry, which were at the time listed on the NYSE. Two sets of time periods were used: one period spanning six years, from 1968-1973, and the second period consisted of two overlapping data periods: 1969-1971 and 1971-1973. The method for evaluating the pollution control was a construction of indices based on two major studies by the Council of Economic Priorities (CEP), conducted in 1970 and 1972. In these two studies, extensive data was collected on 131 pulp mills, owned by the studied firms, for the efficacy of the air and water pollution control systems, which provided the basis for a rating given by the CEP. Financial data and monthly stock price data was provided annually and quarterly by Compustat. The findings of the study indicate that firms in the pulp and paper industry, which has better pollution control, tend to have higher profitability, larger size, lower total and systematic risk, and higher price/earnings ratios than companies with poorer pollution control.
3.4 Summary of Previous Research

3.4.1 Studies on CSR-CFP Relation

Table 2. Summary of Previous CSR-CFP Research.

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Market</th>
<th>CSR Database</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang &amp; Sarkis (2017)</td>
<td>USA</td>
<td>Bloomberg</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Rodriguez-Fernandez (2016)</td>
<td>Spain</td>
<td>Mixed measure from: GRI Index, DJSI, GCN</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Guenster et.al. (2011)</td>
<td>USA</td>
<td>Innovest Strategic Value Advisors</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Filbeck et.al. (2009)</td>
<td>USA</td>
<td>Business Ethics Magazine list of Best Corporate Citizens + KLD</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Brammer et.al. (2006)</td>
<td>UK</td>
<td>EIRIS</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Schnietz &amp; Epstein (2005)</td>
<td>USA</td>
<td>Domini Social Index</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Hillman &amp; Keim (2001)</td>
<td>USA</td>
<td>KLD</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Turban &amp; Greening (1997)</td>
<td>USA</td>
<td>KLD</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>McGuire et.al. (1988)</td>
<td>USA</td>
<td>Fortune 500 Survey on Corporate Reputations</td>
<td>Positive Relation</td>
</tr>
</tbody>
</table>

3.4.2 Studies on CSR-Cost of Capital Relation

Table 3. Summary of Previous CSR-Cost of Capital Research.

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Market</th>
<th>CSR Database</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suto &amp; Takehara (2017)</td>
<td>Japan</td>
<td>Based on Survey by Toyo Keizai Inc.</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Cajias et al. (2014)</td>
<td>USA</td>
<td>KLD</td>
<td>Mixed Result</td>
</tr>
<tr>
<td>Chava (2014)</td>
<td>USA</td>
<td>KLD</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>El Ghoul et.al.(2011)</td>
<td>USA</td>
<td>KLD</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Sharfman &amp; Fernando (2008)</td>
<td>USA</td>
<td>TRI &amp; KLD</td>
<td>Negative Relation</td>
</tr>
</tbody>
</table>
3.4.3 Studies on CSR-Risk Relation

Table 4. Summary of Previous CSR-Risk Research.

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Market</th>
<th>CSR Database</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harjoto et.al. (2017)</td>
<td>USA</td>
<td>KLD</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Nguyen et al. (2015)</td>
<td>USA</td>
<td>KLD</td>
<td>Positive Relation</td>
</tr>
<tr>
<td>Mishra &amp; Modi (2013)</td>
<td>USA</td>
<td>KLD</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Humphrey et.al. (2012)</td>
<td>UK</td>
<td>SAM</td>
<td>No Relation</td>
</tr>
<tr>
<td>Jo &amp; Na (2012)</td>
<td>USA</td>
<td>MCSI (KLD)</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Lee &amp; Faff (2009)</td>
<td>USA</td>
<td>DJSI</td>
<td>Negative Relation</td>
</tr>
<tr>
<td>Luo &amp; Bhattacharya</td>
<td>USA</td>
<td>Fortune Survey + CRSP</td>
<td>Negative Relation</td>
</tr>
<tr>
<td></td>
<td>(2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spicer (1978)</td>
<td>USA</td>
<td>CEP</td>
<td>Negative Relation</td>
</tr>
</tbody>
</table>

3.5 Critique on Previous Research

The measure for volatility used in Lee & Faff (2009), Mishra & Modi (2013) and Harjoto et.al. (2017) was not discussed in those papers, this makes us assume that simple realized standard deviation has been used. Using realized volatility, might not be the most accurate estimate of the true volatility at each point in time. The realized volatility is the observed variation in share prices, however, at each time t there is uncertainty about the price of the next day. This uncertainty factor is not captured by realized volatility. In this study, both simple realized volatility and volatility estimated by a GARCH (1,1) model has been used. This leads to two estimates of volatility. This approach will give this study increased credibility.

The authors of this study would like to critique the CSR data used in previous studies. In previous studies, the CSR data comes from a wide variety of sources, and this study utilizes one additional new source of ESG/CSR performance data. Previously, data has been gathered from the Dow Jones Sustainability Index, as in the case of Lee & Faff (2009). The Kinder, Lydenburg and Domini database or the Domini social index, as in Hillman & Keim (2001), Schnietz & Eppstein (2005) and numerous other studies was the most frequent database used in the past. McGuire et.al. (1988) use the Fortune 500 survey on corporate reputations. Brammer et.al. (2006) uses the Ethical Investment Research Service (EIRIS) in order to gain access to CSR data. These widely different databases all have different methodologies in gathering the data, and Lee & Faff (2009), Filbeck et.al. (2009), McGuire et.al. (1988), Guenster et. al. (2005) and Brammer et.al. (2006) use data
that have a survey as one of the major aspects. Using surveys as the main type of source brings with it a few points to consider. Surveys have a few inherent flaws that may affect the results of the studies above. Low response rate may lead to no-response bias. It can also be difficult to know if the willing respondents are truly representative of the population (Kothari, 1990, p. 101). By using a survey, there is always a risk that the results of the survey are tainted by the respondents individual views of reality. These are often subjective and formed in that individuals reference point and situation. The risk then becomes that the resulting survey is a view into the perception of the respondents, rather than a result that is completely objective and fair. The perception of the respondents then needs to be correct, otherwise the result may not be optimal. It is, however, important to point out that this was not the only source of data in previous studies. Peloza published an article in 2006, stating the general problem of measuring CSR activity objectively. The problem is recognized by McGuire et.al. (1988) as well.

The KLD database is different from the Thomson Reuters ESG Score in a number of areas. According to Chatterij and Levine (2006, p. 41-42) KLD uses absolute bars of performance, instead of relative bars of performance like Thomson Reuters does. What this means is that what KLD does, is to completely rule out certain firms, such as tobacco firms and defense industry companies. In comparison Thomson Reuters give every company an ESG score, regardless of the industry. Chatterij and Levine continue (2006, p. 42) by evaluating the benefits and drawbacks by each of the systems. One downside to relative standards, is claimed to be that ethical standards are often more absolute, for example, ethical investors may not be very interested in the most ethical tobacco company, as they are likely to see all of them as unethical. Some of the benefits of relative bars of performance are, that the “best in industry” firms can be evaluated (Chatterij & Levine, 2006, p. 42-43). Chatterij and Levine (2006, p. 42) furthermore reason around what this would mean, for example if all tobacco companies are excluded from a CSR rating methodology, does this then mean that the companies are under less scrutiny and may thus behave worse? It may. Under a relative standard, the companies are forced to compete against each other, to be “best in industry” (Chatterij and Levine, 2006, p. 42). Using a relative bar of performance helps with giving a more diverse sample, as more types of firms are included in the Thomson Reuters ESG Score. By using the KLD in the previous studies, those studies lack a part of the sample that is there in this study. It may be an unethical part that is lacking, but it doesn’t change the fact that it is part of our shared economy.

The Thomson Reuters ESG score is more appropriate (because of relative bars of performance) and uses both human and artificial intelligence to assure the highest quality of data. While evaluating the studies above, it is evident that the main problem is the quality of the CSR data. This can also be seen in the fact that numerous databases are in use, for CSR data. If one was considered to be the best, surely that would have been used in most, if not all studies. This study is based on the Thomson Reuters ESG score, the methodology that goes into it is rigorous. In addition to this, Thomson Reuters is known for providing high-quality investor services. The Thomson Reuters ESG database has data starting from 2002 (Thomson Reuters, 2018, p. 3). This makes it a relatively new source and previous studies may not have been able to use it due to the relative lack of data history at that time. These arguments lead us, as authors to believe that we are in possession of a dataset that is objective, of higher standard and more unbiased than previous studies in the area. This will be further explored in the Method chapter.
3.6 The Literature Review Process

The literature review process of this thesis started with searching for keywords covering “ESG”, and “volatility”. The following keywords and other combinations of these words were used in order to find articles for the intersection of ESG and the risk-return tradeoff:

- ESG
- CSR
- Sustainability
- Sustainable investment
- Sustainable investment and/effect on CSR
- Sustainable investment and/effect on ESG
- Total Risk and CSR
- Total Risk and ESG
- Volatility and ESG
- Volatility and CSR
- Idiosyncratic risk
- Idiosyncratic risk and ESG
- Idiosyncratic risk and CSR
- CSR and financial performance
- Sustainability and financial performance

References used in the *Theoretical Framework and Literature Review* are scholarly articles. The articles are peer-reviewed and scrutinized. In the chapters *Scientific Method, Method* and *Results* relevant textbooks have been used as references and sources as well. This is done in order to validate for the course of action taken in these chapters and specifically in order to argue for the choices of research theory and statistical methods and tests used. Articles referenced in this study have been accessed through EBSCO (Business Source Premier), JSTOR, Google Scholar, Sage Publications, Springer Publishing, John Wiley & Sons and Emerald Insight. The sources were accessed through a Umeå University library proxy, granting access through agreement.
4. Theoretical Framework

In this part the authors will present the theories relevant to understand the topic being handled, provide a discussion about volatility, as well as introduce the reader to the Life Cycle Sustainability Assessment Methodology. A summary of the key points will be presented at the end of the chapter.

4.1 Modern Portfolio Theory

Modern Portfolio Theory was developed by Harry Markowitz and published in his 1952 paper "Portfolio Selection" in the Journal of Finance. Markowitz was eventually awarded the Nobel prize in economics for its development.

When presented with two portfolios of equal expected return, but with different levels of risk, Modern Portfolio Theory explains that the investor will want to choose the less risky one, i.e. investors are assumed to be risk-averse (Markowitz, 1952, p. 77). Likewise, if two portfolios have the same level of risk, the investor should want to pick the portfolio with the higher expected return. The starting point of this theory is to not view each security's risk and return characteristics as standalone, but as part of a larger context, the total portfolio. Through this it is possible to find the optimal diversification of the portfolio, which can be illustrated through the concept of the efficient frontier in Figure 1. Where R indicates the risk-free asset and the curve indicates the portfolio choices available to investors. E* and σ* indicate the expected return and the risk of the efficient portfolio, respectively (Merton, 1972, p. 1866-1868).

![Figure 1. The Efficient Frontier. Source: Merton (1972, p. 1867)](image-url)
According to Markowitz (1952, p. 79) diversification of a portfolio is achieved by holding instruments that are not perfectly correlated with each other, meaning that a certain expected portfolio return can have different amounts of risk. The risk being diversified away is the security specific risk. The market risk, or portfolio risk, cannot be removed through diversification (Markowitz, 1952, p. 79).

Suppose that you have the required data, which is volatility, expected return and correlations, for all the individual investments considered. Through all of these investment opportunities a set of different portfolios can be created, which all contain their own risk and return profile. You can then select portfolios in two different ways. You can either select all the portfolios that have same risk (volatility), and then choose the one which is set to give the highest return. The other option is to select all the portfolios with the same return, and from these select the one which carries the least risk. The efficient frontier consist of these optimal portfolio choices, for different combinations of risk and return. Portfolios to the right of the Efficient Frontier (in Figure 1) do not have an optimized combination of risk and return, and will because of this not be chosen by a rational investor (Markowitz, 1952, p. 82). There is a rate at which an investor can increase his or her expected return by taking on additional risk, and likewise give up a certain amount of expected return to gain a lower risk (Markowitz, 1952, p. 79). This is a theory which uses a simplified version of reality. One such simplification, is especially interesting in light of this thesis. Namely, the one that states that investors attempt to maximize their economic returns. However, Peloza (2006) states that, recently there has been wide discussion around if investors increasingly also care about aspects such as responsible governance, environmental consideration and ethical business values (Peloza, 2006).

4.2 Capital Asset Pricing Model

The Capital Asset Pricing Model is a simplification and development of Markowitz’s Modern Portfolio Theory. It was presented in the 1964 by William Sharpe, in his article *Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*. This was later further developed by John Lintner in 1965, in his paper *The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets*. The model developed, The Capital Asset Pricing Model (CAPM) is still today widely used in finance. The CAPM shows the connection between level of risk and expected return of an investment. In essence, the theory states that the higher risk which the investor is exposed to, the higher the demanded return should be. The risk premium, reflected in the last part of the formula, is decided by many different factors. Investors opinions of macroeconomic conditions, geopolitical instability or stability, the level of the prevailing and expected interest rate etc. will all have an effect on the perceived risk premium. This will then be reflected in the required returns demanded by investors. The higher the perceived risk premium, the higher the required return will be on any given security, ceteris paribus, this is also visible in the CAPM equation.
The CAPM equation (Fama & French 2004, p. 29) is the following:

\[
\text{Expected Return} = R_f + \beta (R_m - R_f)
\]

Where,

\begin{align*}
R_f & = \text{the risk free interest rate} \\
\beta & = \text{the beta of an individual security} \\
R_m & = \text{the expected return on the market}
\end{align*}

According to the capital asset pricing model, the expected return of a security is the risk free rate, plus the beta of the security times the difference between the return of the market and the risk free rate. The first part of the formula is the risk free rate, which is usually derived from government bonds. For example, U.S. government bonds are considered to be risk free, since the risk of default is essentially zero.

The second part of the formula is the risk premium. By subtracting the risk free rate from the market return, a measure of the risk of the market can be derived. By multiplying the beta of a security with the risk of the overall market, the risk of the individual security will be determined. Fama and French (2004, p. 28-29) state that the asset beta (\(\beta\)) of a security is a measure of how much a security moves relative to the market as a whole. Furthermore, it is a measure of the risk of an investment which cannot be reduced through diversification. To calculate the value of beta, the covariance between the return of the market and the return of the security must be known, as well as the variance of the market returns. Since beta is calculated from historical data, it is exactly that, a measure of what has historically been true for the security, but does not give any guarantee of predicting future performance. It is important to state that there is a difference between expected return and actual return.

To get a sense of what this means, assume a risk free rate of 3% and return of the market of 8%. If the beta is 1, which equals that the security moves exactly together with the market, the expected return would be 3%+1(8%-3%) = 8%. If the beta is 2, the expected return would be 13%.

The capital asset pricing model is relevant to this study because volatility is synonymous with total risk. Fernandez (2015, p. 4) states that the CAPM tells you that the only reason why one stock will, on average, deliver higher returns compared to any other stock, is because it has higher volatility (Fernandez, 2015, p. 4). If a company invests in sustainability, you would certainly want to know if sustainability has an effect on volatility, because according to CAPM you will therefore have altered expected returns. While there are vast amounts of criticism towards this model, it is arguably impossible to exclude from any study examining volatility in a financial context. This is because of its widespread use when dealing with risk and return.

There is criticism on the CAPM as well, as with any simplified theory. Basu (1977, p. 680) finds that low P/E ratio firms generated higher risk-adjusted returns than high P/E returns during the study period, 1957-1971. The author did acknowledge the joint hypothesis problem in the interpretation of the results. Banz (1981) finds evidence for the size effect in data from US stock markets, and argues for that it is an anomaly. Banz does argue that it is still uncertain whether it is a size effect per-se or just an effect that is common among small companies. Jegadeesh (1990) and Fama-French (1992) find that
momentum, i.e. previous performance and firm-level market capitalization and book to market ratio, this implies that returns aren’t only decided by market risk, as prescribed by CAPM.

4.3 Efficient Market Hypothesis and Fundamental Analysis

Starting out in 1953, with a concept that has had and still, to this day has a large impact on finance. The Efficient Markets Hypothesis. In 1953, a British statistician by the name of Maurice Kendall presented a paper on the behavior of stock and commodity prices. Kendall & Hill (1953) had expected to find regular price-cycles in the data, but instead he found that price changes seemed to be random, and not possible to predict by past data. Prices seemed to follow what is called a “random walk” (Kendall & Hill, 1953, p. 11)

Random-walk theorists usually assume that major markets are efficient. Fama & Malkiel (1970) defines an “efficient market” as "A market in which prices always "fully reflect" available information is called "efficient."" (Fama & Malkiel, 1970, p. 383). A “random walk” can be explained in layman's terms as a simple coin-toss game, where each outcome of a new coin-toss is independent from the previous. For example, a heads-outcome on the first toss, does not make a heads-outcome any more likely on the second toss. Similarly goes the logic for random walks. Previous price changes cannot be used to predict future price changes, the two are independent (Kendall & Hill, 1953, p. 11)

There are generally two predictive techniques used by professionals in order to try to beat the market, the theory of technical analysis and the theory of fundamental analysis. Fama (1965, p. 39) states that what makes subsequent price changes independent, may be the existance of sophisticated traders, both chart readers i.e. technical analysts and fundamental analysts. The authors of this paper expect investors concerned about ESG-factors to generally belong to the fundamental analysts, basing their decisions on fundamentals. Fundamental analysts, believe that each share has an intrinsic value. Intrinsic value means that every share on the stock market, and by extension of course every underlying company, has a value that is ‘fair’. This fair value is based on the assets and the liabilities of the company in question, as well as some discounted value of future earnings prospects and intangibles (non-material assets, such as brand, patents etc.). Fama (1965, p. 38-41) states that by studying these aspects, analysts should in theory be able to find the intrinsic value and then compare it to the market price, and thus determine if a given security is either under- or overpriced. This is done by evaluating new information that reaches the market, having a superior method of analysis may still benefit investors. However, Fama (1965, p. 39-40) continues with adding that when there are enough superior intrinsic value analysts in the market, prices will reflect the new information that hits the market very quickly. This does not imply that superior intrinsic value analysts cannot make larger profits than investors that use a simple buy and hold strategy, but this depends on consistently being able to predict and evaluate new information accurately. Evaluating this information too late, yields no benefit, as presented earlier, the presence of sophisticated traders play a role in keeping price changes independent from each other, as the findings by Kendall & Hill (1953) showed that they are (Fama, 1965, p. 38-41).

Why is this interesting in light of this paper? Fundamental analysis is a very broad term and encompasses analysis on almost any variable that is produced by a company. Similarly as with technical analysis, the data is unbiased before investors analyse the data. As the analysis is carried out the fair value of the share analyzed becomes biased. This
leads to different views on the fair price on the analyzed share, which then leads to volatility. What is interesting about this process is the inputs used by the investors. “Fundamentals” in the traditional sense would mean a focus on profitability, dividend ratios, change in dividend, change in profit, book to market values and price/earnings ratios et cetera. However ESG-investors may be interested in a new type of fundamentals, that are more and more widely reported. These are, as mentioned before, factors concerning environmental, social and governance aspects. However many investors still use price / earnings ratios (P/E), different type of book values and a diverse span of profitability measures, both accounting based, such as net margin and also market based ones, such as P/E.

Fama & Malkiel (1970, p. 383) define three different forms of market efficiency. First is weak-form market efficiency. It states that market prices of securities today, contain all information about historical prices. This would have a couple implications. First, technical analysis would be useless. Since all available information about the prices of securities is already priced in the market. Furthermore, prices would follow a “random walk”, as identified by Kendall & Hill (1953) Second, is the semi-strong form of market efficiency. Furthermore, Fama & Malkiel (1970, p. 383) state that the semi-strong form includes not only all information about historical prices, but also all publicly available information. This includes annual reports, financial press and basically all of the information contained on the internet. This implies that any new information released to the market, would immediately be reflected in stock prices. For example, upon release of a new quarterly report, the market would immediately respond to the new information available and this would be reflected in the share price, without delay. This would make fundamental analysis useless, as well. Finally, is the theory of strong form efficiency. Strong-form market efficiency states that market prices reflect all available information. This includes possible inside information. It implies that investors would only be either lucky or unlucky, and no skill would be involved in stock-picking. Fama & Malkiel (1970, p. 383) concludes that apart from certain instances, the efficient market hypothesis holds up well.

In spite of this, there is evidence both for and against efficiency, the issue is debated still today and no clear consensus has been found. On the one hand, the evidence against efficiency, there are a number of anomalies found that conventional asset pricing methods cannot explain. An anomaly is defined as any return that isn’t captured by the capital asset pricing model, which is the usual tool used when calculating expected returns. In a study of long-term anomalies De Bondt and Thaler (1985) state that investors put too much emphasis on past performance and too little emphasis on the fact that performance tends to be mean-reverting over time. These reversals are accounted for by the argument that investors overreact to past performance. This would be an argument for contrarian investing, i.e. investing in firms with weak current performance would give abnormal returns, since focus is put on past performance by investors. An argument against efficiency. Lakonishok et al. (1994) make a similar finding, adding that the market seems surprised when returns mean-revert. Ritter (1991) and Loughran and Ritter (1995) find that the post-event returns of IPOs are poor in the long-term, in other words they underperform. The momentum effect is another well known anomaly, found by Jegadeesh and Titman (1993). Shares with high returns over the previous three to six months, tend to have high return over the coming three to six months. This short snapshot of anomalies proves that there is an ongoing discussion in this matter and it isn’t resolved yet.
Furthermore, there is widespread criticism and contrary evidence about the viability of both the Efficient Markets Hypothesis and the Capital Assets Pricing Model. Fama states in his 1991 article *Efficient Capital Markets: II*, that the major problem with disproving both the theories is the joint hypothesis problem. For example, small companies are sometimes claimed to have higher returns than large companies even after adjusting for beta. One hypothesis is the efficient markets theory, the second hypothesis is the CAPM. In the example mentioned, it is not possible to distinguish which theory fails. In order to distinguish abnormal returns, an asset pricing model needs to be used. So, any inefficiency can be interpreted as an anomaly, faulty asset pricing model or both (Fama, 1991, p. 1575-1576)

4.4 The Portfolio Characteristics of Individual Investors

In reality, the portfolios of investors have a few distinct characteristics, that are not entirely consistent with Modern Portfolio Theory. By looking at brokerage accounts, Barber & Odean (2000) show that investor trading seems to be excessive. This conclusion is reached by comparing gross and net return. The reduction in net returns is due to trading. Even after removing most trades that are likely to be motivated by tax-loss selling, portfolio rebalancing, etc. Trading still lowers returns. This is of course due to transaction costs. Goetzmann & Kumar (2008) show that private investors are underdiversified, i.e. they hold fewer stocks than normative theory would suggest. The dataset is similar to the one used by Barber & Odean (2000), a large online brokerage. Polkovnichenko (2005), makes similar findings, investors do hold too much idiosyncratic risk in their portfolios, according to the findings of the author. The conclusion of Goetzmann and Kumar (2008) is that the unexpectedly high idiosyncratic risk results in a welfare loss. It is also argued that most individual investors would have been better off just investing in a passive index fund. French and Poterba (1991), Grinblatt & Keloharju (2001), Kilkka & Weber (2000) all find that domestic stock ownership is relatively too high and that investors prefer to own stocks from domestic rather than foreign markets, this is generally known as “home bias”. This means that the resulting portfolio is not an efficient one, it is possible to earn a higher return at the same level of risk, or lower the risk and still obtain the same level of return, if investors would diversify abroad, as normative theory suggests.

4.5 Volatility

Risk in investments, means that future returns are unpredictable. This unpredictability is measured using standard deviation or variance. In the context of the stock market, two broad types of risks are generally identified. Systematic- and idiosyncratic risk. The two are components of total risk, or volatility. Systematic risk is defined as risk that cannot be avoided, assuming participation in the stock market. This is because there are certain risk elements that are market-wide, these risks cannot be avoided by portfolio diversification. However, most investors don’t only own shares in one company, they hold a portfolio of shares. Idiosyncratic- or firm-specific risk is defined as risk that can be “diversified away”, by holding a diverse portfolio. Stocks, while highly correlated between one another, are not perfectly correlated. This is what allows for the firm-specific risk to be “diversified away”, diversification reduces the variability of returns. According to theory, an investor should keep diversifying until only the market risk remains (Markowitz, 1952). So, according to traditional theory the systematic risk is the only priced risk-factor
in the market. Recent empirical findings, however, suggest that it might not be quite this simple.

What aspects do or don’t cause volatility? Shiller (1981) finds that price reaction to changes in dividends can be considered to be “too big”. Furthermore, LeRoy and Porter (1981) find that using the present-value equation, which states that the coefficient of dispersion of share prices must be lower than that of dividends. They found that volatility was higher than would be predicted by the efficient market hypothesis. Roll (1988) finds that it is not easy to explain what causes share price changes, even after the fact. In his study he finds that systemic economic effects, other firms shares return in the same industry and public firm-specific news announcements have relatively little explanatory power. The study found a monthly R-squared of only 0.35 and a daily of 0.20. These studies show that fundamentals alone cannot explain share price fluctuation.

DeLong et al. (1990) claim that noise traders have extremely wide and unpredictable expectations. Arbitrageurs exercise caution because of this uncertainty. This means that prices can diverge significantly from their intrinsic values, even if there is no change in the underlying fundamental risk. Campbell and Kyle (1993) create a model where the stock price is equal to expected dividends discounted at the risk-free interest rate, minus a constant risk premium of the market, plus a noise trader aspect. They find, just as DeLong, et. al that noise trading increases volatility. Black (1986) as paraphrased by DeLong et.al. (1990, p. 704) defines a noise trader as a trader that does not have inside information, but rather just acts on “noise” (news) as if this information would give them an advantage. This is suggestive to that volatility arises partly from the process of trading, without any fundamental view of the market while doing so.

Moreover, Campbell et. al. (2001) find that the idiosyncratic risk has been increasing relative to market risk in recent years prior to that piece of research, by studying individual shares on the market, industry and firm level. This implies less explanatory power of traditional market models, such as CAPM. Wei and Zhang (2006) follow up this study, and find that firm earnings have decreased and become more volatile during the same time period as in Campbell et. al. Furthermore, volatility was found to be negatively correlated with return-on-equity, and positively related with the volatility of return-on-equity in the cross-section. On the other hand, Brandt et. al. (2010) finds that the spike in volatility found by Campbell et. al. (2001) and Wei & Zhang (2006) was episodic and by 2003 volatility had decreased to pre-1990s levels. This points to the nature of volatility. Characterized by episodes of high volatility, and subsequent periods of lower volatility. High volatility tends to follow high volatility, and vice-versa.

Kumar & Lee (2006) find that, by controlling for retail investor concentration, it is evident that firms with high arbitrage cost, and higher idiosyncratic risk, show much stronger sensitivity to changes in retail investor sentiment. Ang et.al. (2006) find that firms with high idiosyncratic risk have very low returns, even negative in their sample. The effect cannot be explained by size, book-to-market, leverage, liquidity, volume, turnover, bid-to-ask spread, coskewness or the range of analysts’ forecasts. The result is very robust and persists in both bull and bear markets. It is also argued that total risk i.e. volatility is a new systemic factor (Ang et.al., 2006). These papers indicate that idiosyncratic risk matters. However, market-wide risk factors will likely remain even in the event that volatility would prove to be lower. This would then imply a lower firm-specific risk component.
Malkiel & Xu (2003) study the behavior of idiosyncratic volatility after the World War II period. The approach is different from Campbell et.al (2001). Aggregate idiosyncratic volatility statistics are constructed from the Fama-French three-factor model. The Fama-French three factor model is an extension on the CAPM, it includes factors such as size, book-to-market, as well as the market risk premium, similar to CAPM (Malkiel & Xu, 2003). By using this approach, Malkiel & Xu (2003) find that idiosyncratic risk seems to have increased over time. Strong evidence is found that idiosyncratic volatility is positively related to institutional ownership, even after controlling for size effects. Fu (2009) finds evidence supporting this view. EGARCH models are used to estimate expected idiosyncratic risk. It is found that idiosyncratic volatility is positively related to expected returns. Again, this supports the view that investors are undiversified and that idiosyncratic risk does play a role for investors.

**4.6 Life Cycle Sustainability Assessment Methodology**

The Life Cycle Sustainability Assessment Methodology (LCSA) is a concept that has developed from the German “Produktlinienanalyse”. The concept was put into a formula by Kloepffer (2008), (As cited by Traverso et. al., 2010 p. 3312). The formula is the following:

\[
\text{LCSA} = \text{LCA} + \text{LCC} + \text{SLCA}
\]

\[
\text{LCA} = \text{Environmental Life Cycle Assessment}
\]

\[
\text{LCC} = \text{LCA-type Life Cycle Costing}
\]

\[
\text{SLCA} = \text{Social Life Cycle Assessment}
\]

Kloepffer (2008, p. 90-93) states that LCA includes within it the environmental part of sustainability. Using this framework can bring several different benefits:

- **Life Cycle Perspective**, this component includes the management of the entire life cycle of a product. From raw material to finished product. This takes its form as a more complete planning of all aspects of a business, using recycling and reusing of old materials as a way to lower cost and decrease product life-cycle cost, which is a both financial and environmental cost, this approach was also described to be in use by the CEO of Dow Chemical in a Harvard Business Review interview (Harward Business Review, 2017).
- Helping to achieve strategic business aims by incorporating environmental issues into the management of the business.
- Giving a solid foundation to develop competitive advantage, by improved efficiency and reducing costs.
- Encouraging environmental performance of suppliers, by integrating them into the business systems of the organization.

Kloepffer (2008, p. 90-93) furthermore states that, while LCA takes the perspective of the environment and how to manage it alongside the business, LCC takes the perspective of economic matters and how to manage LCA profitably. Figure 2 provides an oversight into the crucial aspects of each. As can be seen, on some of the points the differences are such that they are difficult to manage simultaneously and one might have to be prioritized over the other.
For example, the treatment of time and scope. Timing cannot be simultaneously ignored and taken into account. This difference in the scope of planning could potentially cause problems. Norris (2001) in his paper “Integrating Economic Analysis into LCA” argues that this gap between LCA and LCC leads to a series of problems. First, companies cannot afford to make decisions solely on the basis of LCA. Product performance, cost as well as the cost of alternatives need to be weighed in, also considering our limited resources. Second, the separation of LCA and LCC leads to missed opportunities in leveraging both types under the same decision making framework. Leveraging both could allow for optimal decisions in questions of product design. The question asked would then be along the lines of, how can we design this product in a way that is environmentally sound and low cost? In other words, both LCA and LCC should be integrated into each other to be able to make the optimal decision. Third, the LCA perspective can have economically sound results when applied correctly within LCC. For example, economic analysis of factors such as pollution, can lower the cost when excess pollution is avoided. This is even more so the case when the avoided risks of accidents, reputation and the environment is taken into account (Norris, 2001, p. 61). Peloza (2006) supports this view in his paper stating that sounds practises in CSR, which is closely linked to this issue, works as a form of insurance. Cost of buying insurance might be lower, since the risk for accidents might be lowered (Peloza, 2006).

The LCA and SLCA frameworks include the same aspects that the Thomson Reuters ESG score aim to measure. SLCA is concerned with issues within the dimensions of Human Rights, Labour Practises, Society and Product Responsibility. Jørgensen et.al. (2008) recognize that this aspect of the LCSA framework is a new and immature field, this becomes evident in the apparent trouble to measure the impact of the efforts made by companies. There are, at the moment both quantitative and qualitative approaches to the issue. I.e. the measurement problem also evident in the literature review of this paper is emphasized. There is also an ongoing debate into what aspects to include under the SLCA
umbrella. Another issue is that the impact of social work is very much a subjective one and the success or failure partially depends on the reference point (Jörgensen et.al., 2008).

A framework such as this one, can be viewed as a rough guideline that helps companies in getting for example ISO-certified. This serves as a tangible proof of the sustainability work being done. Amazon, has at least parts of its supply-chain certified by some kind of ISO certification (Amazon, 2018). An ISO commissioned survey found that over 1 million new certificates were awarded in 2014, this proves how common it is for companies to use some kind of quality assessment framework (ISO, 2016).

4.7 Summary of Theoretical Framework

**Modern Portfolio Theory**
- A rational investor strives to maximize return, while minimizing the risk carried in investments.
- The amount of risk held can be lowered by diversifying.
- Modern Portfolio Theory has a focus on how an investor should behave.
- Modern Portfolio Theory assumes that systematic risk is the only priced risk factor.

**Capital Asset Pricing Model**
- The CAPM is concerned with the risk-return tradeoff.
- States that the higher the systematic risk or asset beta, the higher return should be expected by an investor.
- Systematic risk is the only priced risk factor in the market.

**Efficient Market Hypothesis and Fundamental Analysis**
- States that prices follow a “random walk”, i.e. price changes are independent of each other.
- An “efficient market” is defined as “A market in which prices always "fully reflect" available information” (Fama, 1970, p. 383).
- Three forms of market efficiency are generally identified. The weak-, semi-strong-, and strong form of market efficiency.
- Fundamental analysis is the focus by investors on company fundamentals in the analysis of share prices.

**The Portfolio Characteristics of Individual Investors**
- Individual investors trade too much.
- Individual investors are underdiversified.
- Individual investors are prone to “home bias”
- It is argued that these characteristics lead to individual investors holding portfolios that are not the efficient portfolio.

**Volatility**
- Total risk i.e. volatility, consists of systematic and idiosyncratic risk.
- Previous research suggests that fundamentals alone cannot explain the volatility of share prices.
- Previous research suggests that noise traders can affect market prices and volatility arises partly from the process of trading.
• Previous research suggests that volatility is episodic. Periods of low volatility are followed by periods of low volatility, while periods of high volatility are followed by periods of high volatility.
• Previous research suggests that idiosyncratic risk might matter to investors.

**Life Cycle Sustainability Assessment Methodology**

• The LCSA framework provides a possible template for companies to base and evaluate their CSR efforts.
• Managing the aspects of LCSA well, should allow companies to integrate aspects of CSR into their company, while also taking profitability into account while doing so.

The theories presented in this chapter has introduced the reader to how investors view risk, and how traditional theory reasons around the issue. The LCSA framework is introduced, in order to lay a foundation as to what aspects are troublesome when trying to manage CSR within companies, how the aspects of CSR can relate and offer some possible solutions to some issues. The *Theoretical Framework* is meant to make analysis of results more fluent and deep, in conjunction with the *Literature Review* presented. The rationale for including these theories goes along the same lines as outlined in section 3, they are deemed most appropriate.
5. Method

This part will present the different aspects of practical method used in the thesis. First, details about the sample will be provided, as well as a presentation of the ESG score methodology. Second, the independent variables used will be explained in detail. Third, the statistical method will be explained.

5.1 Population and Selection of Sample

The quality of a piece of research is affected among other things by the sample, and how it is selected. Cohen et. al. (2011) urges researchers to make the decision on how to choose the sample early in the research process because it allows for planning, and better time-management (Cohen et.al., 2011 p. 143). A sample is defined as the subset of data that is analyzed in a study, while a population is the entire group to which the study aims to generalize the results (Bryman & Bell, 2015 p. 191). In this study, the sample that has been chosen by the authors consists of companies in the S&P 500 index. The S&P 500 is divided into subdivisions according to the Global Industry Classification Standard (GICS), these are effectively clusters of firms divided into groups according to their line of business. To be more specific, the index is subdivided into 11 sectors, 24 industry groups, 68 industries and 157 sub-industries. The 11 sectors represent every possible line of business among them. The authors of this paper have attempted to select all of the companies on the S&P 500 index. Due to the lack of complete data, 23 companies have been deleted from the sample. This leaves 481 companies, which were included in the sample. For a list of deleted companies, see Appendix 1. By exactly identifying the missing companies, barriers to making replicating studies is minimized, as there is no ambiguity of which companies are included or excluded in the sample. The scope of this study is between the January 1st 2009 and December 31st 2016. This gives us 2014 data points of daily closing prices of the companies included. The population onto which generalizations are attempted to be made, is the world-wide stock markets. There are also a number of ESG Scores missing in the sample. The reason for this is unknown, but assumed to be a lack of reporting, either from the companies themselves or on Thomson Reuters end, the reason is not viewed to be relevant enough to warrant for time being spent on finding out, since the amount is negligible and will not affect the results of this study.

5.2 Thomson Reuters ESG Score

The data used in this study was collected from the Thomson Reuters Eikon-platform, and the data is collected and pre-analyzed by Thomson Reuters. If this dataset was not available, the authors would have been forced to collect the data themselves, by studying financial reports, media reports, press releases, making on-site visits and create a mechanism to classify and analyze all the raw data collected. This would certainly have been impossible, both because of time constraints and costs as well as lack of expertise. The dataset allows for a study to be done that otherwise would have been impossible. Access to the platform was provided for this study by the Umeå University library. Thomson Reuters is a world leading source of news and information for professional
markets. As this is one of the major parts that may affect our results in various ways, a discussion about the dataset, methodology behind it and what it entails is in order.

Thomson Reuters (2018) recognize that many investors today strive to invest responsibly. The company aims to support that by giving out these ESG scores. The company states that the “Thomson Reuters ESG Scores are designed to transparently and objectively measure a company’s relative ESG performance, commitment and effectiveness across 10 main themes (emissions, environmental product innovation, human rights, shareholders, etc.) based on company-reported data.” (Thomson Reuters, 2018, p. 3). The scores themselves are available in simple percentile rank form. They are available both in percentage (i.e. scale 0-100) and as a letter grade from D- to A+. This study will use the percentage measurement. These scores are benchmarked against Thomson Reuters Business Classifications (TRBC) Industry Group. In other words, they are benchmarked per industry group, making comparison across industries possible. We, the authors of this paper, have interpreted this as there being no added value in including an industry dummy in this study, since this distinction is made on the raw-data level. The ESG scores are an enhancement and improvement upon the Asset4 database (Thomson Reuters, 2018, p. 3). Comparability is enhanced, by introducing benchmarks on the industry and country levels at the data point scoring level. Data driven category weights are added, in order to reflect the data availability within each category, supporting differentiation across companies. A percentile rank scoring methodology was added, to eliminate hidden layers of calculations (Thomson Reuters, 2018, p. 3).

Figure 3. ESG Score Overview. Source: Thomson Reuters, 2018.

Figure 3 gives an overview of what the Thomson Reuters ESG score includes. The ESG combined score and the controversy score have been omitted in this study. The reason for this is that the controversy score is only altered in years when there is a scandal of some sort, this implies that in years where a scandal is absent, the controversy score would remain on the same level both for a company with high ESG score and low ESG score. From this, it was concluded that the measure is not good enough to be included in this study. As the combined score is an aggregate score of both the ESG score and the controversy score, this also rules out the use of the combined score. Breaking down the ESG score, the environmental aspects investigated by Thomson Reuters are; (1) Resource use, this includes the ability of the company to use as little as possible in terms of materials, water and energy. Finding more eco-friendly approaches is seen as a benefit. (2) Emissions. This category measures a company’s effectiveness and commitment in
reducing emission in production and other operational processes. (3) Innovation. It reflects the ability of the company to innovate in a way that reduces environmental costs and burdens for its customer base, this can be done by creating new market opportunities, environmental technologies or designing eco-friendly products (Thomson Reuters, 2018, p. 15).

The social aspects include the following; (1) Workforce. It includes the effectiveness in promoting job satisfaction, a safe and healthy workplace, being open to diversity and supporting equal opportunity, as well as development opportunities. (2) Human rights. Simply investigates how effective a company is in upholding fundamental human rights conventions. (3) Community score. Measures the commitment of a company in adhering to being a good corporate citizen, protecting public health and conducting ethical business. Lastly, (4) the product responsibility score. It measures the ability to produce quality goods and services. While upholding the health, safety, integrity and data privacy of the customer (Thomson Reuters, 2018, p. 15).

The governance aspects include; (1) Management. Measures the commitment and effectiveness in following best practice corporate governance principles. (2) Shareholders score. Measures the equal treatment of shareholders and use of anti-takeover defenses. (3) The CSR strategy score, reflects the company’s efforts in integrating economic, social and environmental aspects into the day-to-day processes (Thomson Reuters, 2018, p. 15).

Furthermore, Thomson Reuters (2018, p. 6-8) state that each of these ten minor categories, have within them multiple points that are being evaluated. For example, the workforce subscore, contains of 29 different measures. CSR Strategy score has the smallest number of individual subcategories, with 8. In total there are over 150 ESG indicators that are collected and analyzed before subsequently forming the ESG score. For an overview of the weights that each of these categories hold in the ESG Score, see Appendix 4. This can be seen as quite an extensive effort into gaining knowledge into how well a company actually manages ESG aspects. By studying company reporting, it should through these measures, be revealed what companies perform well in reality and what companies only engage in greenwashing. “Greenwashing is the selective disclosure of positive information without full disclosure of negative information so as to create an overly positive corporate image (Lyon & Maxwell, 2011)”, as cited by Bowen and Arragon-Correa (2014, p. 107).

The data collection process is of essence when compiling a database of any kind, and the reliability of the final result is highly affected by the processes used (Kumar, 2011, p. 902). This is why Thomson Reuters describe their process at length. The data is collected from six major sources, these are; annual reports, company websites, NGO websites, stock exchange filings, CSR reports and news sources. The data is collected by over 150 research analysts worldwide, trained to collect ESG data. To achieve the best possible data quality, both human and algorithmic processes are used. Four phases in the process are identified, the initial data entry process has over 400 built in error checks. The post-production includes 300 different quality screeners on many aspects of the data. Independent audits are carried out. Lastly, management review the data and process (Thomson Reuters, 2018, p. 15).

Kumar (2011, p. 450-451) states that using secondary data comes with a few pitfalls. Validity and reliability should be scrutinized. The data collection process described above
is thorough enough to reasonably think that validity and reliability of the data should be high, the ESG scores are likely to measure what they are meant to measure. The validity and reliability of this study as a whole will be analyzed later. Kumar (2011, p. 450-451) carries on, stating that bias of the data collector is another risk with secondary data, in this case, many different people and even artificial intelligence is used to minimize the risk of this issue (Kumar, p. 450-451).

5.3 Data Management and Control Variables

In this section the process of how the data has been handled and how the control variables were created is described, as well as giving the reasons for including each control variable. The authors of this paper spent time searching for sources for stock closing prices, earnings per share data, debt level data etc. Different sources were considered, such as Morningstar and Yahoo Finance etc.. It became evident that Thomson Reuters Eikon provided the data that was needed. As a result, all data for this study has been gathered from the database. Many approaches to data collection can be both time-consuming and expensive, especially when original data is to be collected. Many institutions and companies collect data on behalf of investors or as part of their regular business (Bryman & Bell, 2015, p. 307). This study utilizes such secondary data.

5.3.1 Volume

Volume in this thesis indicates trading volume, that is how many shares have changed owners during each specific day. It might be reasonable and intuitive to reason that volatility arises from new information reaching the market and market participants reacting to this information, by trading. This is supported in research. For example, Fama (1965) find that the variance for a trading period from Friday (closing) until Monday (closing) is only 22% more than the average variance on one trading-day during the week. The authors of that paper conclude that it reasonably should be almost three times the one day variance midweek. This implies that some of the variance arises from the mere possibility to trade (i.e. people trade because the market is open). The alternative hypothesis in that study was that people trade three times as much on Monday, due to a building-up of new information over the weekend. Furthermore, French and Roll (1986) find that asset prices are much more volatile during exchange trading hours than during other times. The study finds three reasons for it, first, volatility is caused by public information being more likely to be released during regular business hours. Second, volatility is caused by private information which affects prices when informed investors trade. Third, volatility is caused by pricing errors caused during trading. These studies give support to including volume as a control variable, as it has been suggested that it might affect volatility by previous research.

5.3.2 Market Capitalization

This variable is used to measure the size of the companies. The authors were able to retrieve data on weighted average shares outstanding during each year as well as the share price, these were then multiplied to get market capitalization on a daily level. It is reasonable to believe that the size of the firm may affect volatility. Perez-Quiros & Timmermann (2000) point to the asymmetries in risk between large and small corporations across the business cycle. As suspected, small firms do show a higher risk according to their measure than large corporations. This is evident both in recession and
expansion, and this is translating into the share prices of the companies according to Perez-Quiros & Timmermann (2000). Below is the formula used in order to receive the variable Market Capitalization:

\[
\text{Daily Share Closing Price} \times \text{Yearly Weighted Average Shares Outstanding} = \text{Daily Market Capitalization}
\]

5.3.3 Total Debt to Enterprise Value

Total Debt to Enterprise Value is included as a measure of the debt level of each firm. This variable is retrieved in its unedited form. The ratio compares the total debt load to the value of a firm’s debt and equity. One could assume that debt has an impact on volatility, at least when the debt becomes troublesome for a company to manage. Fama & French (2002, p. 1) state that the trade-off theory of capital structure indicates that the firm is perfectly leveraged when the benefit of one more dollar of debt is exactly outweighed by the cost of taking on that dollar of debt. The costs of high debt include potential costs of bankruptcy as well as agency conflicts between stock- and bondholders. Based on the trade-off theory and Fama & French (2000) it is reasonable to believe that such problems would have an impact on the investors expectations of the share price fluctuation.

5.3.4 Price / Earnings Ratio Including Extraordinary Items

The Price / Earnings ratio (P/E ratio) is and has been one of the most commonly used ratios in company valuation. In this study, the ratio captures the current share price, divided by the most recent fiscal quarter’s earnings per share (EPS) figure. According to Shen (2000, p. 24) A relatively high P/E indicates that future stock prices might decline. However, a negative P/E ratio means that a company is making losses in the current period. A low P/E ratio might indicate that prices are about to increase in the future. Shen (2000, p. 24-25) goes on to argue that P/E ratios have a tendency to revert towards their long-run mean. It is deemed reasonable to expect companies with a low P/E ratio to be expected to have a higher volatility, since the earnings prospects might be more uncertain and investors doubts about the company might increase and in this study low P/E ratios include companies with negative current earnings. The following formula was used to get the P/E ratio.

\[
\frac{\text{Daily Share Closing Price}}{\text{Yearly EPS Including Extraordinary Items}} = \text{Daily P/E Ratio}
\]

5.3.5 S&P 500 Index Return

The S&P 500 index has been included as a control variable to compare the volatility changes in individual firms to average market returns. Logarithmic returns have been used. The formula for obtaining this variable is the following:

\[
R_t = \ln \frac{P_t}{P_{t-1}}
\]
5.3.6 S&P 500 Realized Volatility

The realized volatility of the S&P 500 Index has been included to control for the overall volatility of the market during the timeframe of the study. The formula used in section 5.3.5 provides us with logarithmic returns, in order to convert this to Realized Volatility, the following formula was used:

\[ R_t^2 = \sigma^2 \]

5.4 Statistical Analysis

In this thesis, two different methods of measuring volatility have been used. Realized volatility and a GARCH (1,1) estimate of volatility. The realized volatility has been obtained by using the following formulas:

\[ R_t = \ln \frac{P_t}{P_{t-1}} \]

This equation was used in order to get daily logarithmic returns, for each of the companies in the data set, for each individual day in the sample. In order to receive a estimate of the daily volatility, the daily logarithmic returns were squared.

\[ R_t^2 = \sigma^2 \]

\( \sigma^2 \) is the proxy used for realized daily volatility and subsequently one of the dependant variables in the study. The formulas were used on each company separately in Microsoft Excel, in order to receive a daily volatility proxy for every company included in the sample.

The GARCH (1,1) estimate for daily volatility, that was subsequently used as a second proxy for volatility in this study, has been obtained by the help of the Stata 15 software. The command used in order to estimate GARCH (1,1) volatility in Stata is “arch return if companyid ==x, arch(1/1)garch(1/1)” where, “return” is retrieved by using the formula for logarithmic returns “companyid ==x” indicates the companies 1 to 481 included in the study, one number assigned for each company. This command was thus used for each company individually. The equation is the following:

\[ \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta \sigma_{t-1}^2 \]

Where, the first term, \( \sigma_t^2 \), is the conditional variance term. The conditional variance is allowed to conditionally depend on the variance in the previous period, indicated by \( \alpha_1 u_{t-1}^2 \) in the equation and the fitted variance from the previous period, indicated by \( \beta \sigma_{t-1}^2 \) (Brooks, 2014, p.428). Subsequently after this, the post-estimation command “predict Var_X if companyid ==x, variance” was used for every company after using the first command above. By inserting these commands, the estimated conditional variance was retrieved for each company.
5.4.1 Conditional and Unconditional Variance

In the previous section, two different variance estimation models have been presented, that will be used as proxies for the volatility of the companies shares. Brooks (2014, p.424) states that the unconditional variance is synonymous to the simplest form of realized volatility estimate. The variance is unconditional, because we are not conditioning it on the information available today. On the other hand, conditional volatility is derived from the GARCH (1,1) prediction of conditional variance. It is conditional upon the information we have today, as shown by the GARCH (1,1) equation.

5.4.2 Panel Data Analysis

Panel data is defined as “the pooling of observations on a cross-section of households, countries, firms, etc. over several time periods” (Baltagi, 2005, p. 1). In this case, observations are firms and the time period is measured in days. 2014 is the maximum day number included in the study. Compared to data which uses only a cross sectional or time series setup, panel data allows identification of certain questions or parameters, without the need to implement restrictive assumptions. This would express itself in the possibility to investigate changes which occur on an individual level. For example, assume that sales of bicycles increase by 5% from one decade to the next. With panel data, it is possible to answer if this change is due to an increase of 5% for all individuals, or 10% increase for half of the individuals and no change for the other half. Panel data is not only suitable to explain why individual units behave differently, but also why a given unit behaves differently at different time periods (Verbeek, 2012, p. 373). This method was deemed suitable in this study, since that is a prerequisite to answer the research question that has been proposed. Panel data accounts for heterogeneity within companies (Baltagi, 2005, p. 4-5). Baltagi (2005, p. 4-5) continues with stating that using panel data gives more informative data and less collinearity between variables. Furthermore, Baltagi (2005, p. 11) states that a panel data model is different from a simple time-series model in that it has a double subscript on the variables in the equation, this indicates the design of it, the i stands for the cross-sectional dimension and the t for time.

5.4.3 Regression Models Used in the Study

The regression model used in this study will be a panel regression model. The regression model will have two separate measures of volatility. Realized Volatility and a GARCH (1,1) estimate of volatility but the model is otherwise identical for both methods of volatility estimation. The equation below is the regression model used in this study.

\[
\hat{\sigma}_{it} = \alpha + \beta_1 ESG\ Score_{it} + \beta_2 Volume_{it} + \beta_3 Market\ Capitalization_{it} + \beta_4 Debt\ to\ Enterprise\ Value_{it} + \beta_5 P/E\ Including\ Extraordinary\ Items_{it} + \beta_6 S&P\ 500\ Return_{it} + \beta_7 S&P\ 500\ Realized\ Volatility_{it} + u_{it} \hat{\sigma}_{it}
\]

Where, \(\alpha = \text{intercept}\)

\(\beta_1 X_{it} = \text{Coefficient and independent variable}\)

\(u = \text{error term for the dependent variable}\)
5.4.4 Fixed versus Random Effects

Fixed Effects

What a fixed effects model does is it essentially focuses on differences between the ‘within’ individuals of the data (Brooks, 2014 p. 531). The model investigates the relationship between outcome and predictor variables within an entity, for example a company or person. In the case of this study, this is the different companies used (coded by companyid 1, 2, 3...504 in the panel data). Each of these entities has its own characteristics which can influence the predictor variables. The fixed effects model is suitable when it is suspected that something within the individual may bias or impact the outcome or predictor variables. These are characteristics that do not depend on time, for example being a man or a woman, and needs to be controlled for, which is done through the fixed effects model. This will give a corrected result of the effect from the variables used in the model on the dependent variable. It is commonly assumed that all \( x_{it} \) are independent of all \( u_{it} \) in the fixed effects model (Verbeek, 2012, p. 377-379). The fixed effects model is written as:

\[
\hat{\sigma}_{it} = \alpha_i + x_{it}'\beta + u_{it} \quad \text{where } u_{it} \sim \text{IID} (0, \sigma_u^2)
\]

- \( \alpha_i \) (i=1...n) is the unknown intercept for each entity (n entity-specific intercepts).
- \( \sigma_{it} \) is the dependent variable where i = entity and t = time
- \( x_{it}' \) represents one independent variable
- \( \beta \) is the coefficient for that independent variable
- \( u_{it} \) is the error term (Verbeek, 2012, p. 377-379)

Random Effects

It is commonly assumed in regression analysis that all the effects on the dependent variable, which can not be explained by the regressors in the model, can be summarized by a random error term (Verbeek, 2012, p. 381). Alpha (\( \alpha \)) thus represents random factors, independently and identically distributed over individuals. Together with the term u, an error term is formed (\( \alpha + u \)) which accounts for changes which do not vary over time, and a remainder component, which is assumed to be uncorrelated over time. The random effects model is written as (Verbeek, 2012, p. 381).

\[
\hat{\sigma}_{it} = \beta_0 + x_{it}' + \alpha_i + u_{it} \quad \text{where } u_{it} \sim \text{IID} (0, \sigma_u^2); \alpha_i \sim \text{IID} (0, \sigma_u^2)
\]

Difference across entities having influence on the dependent variable is the reason for using the random effects model. In the random effects model you can include time invariant variables, which is not the case in the fixed effects model. Examples of this can be gender or birthplace (Verbeek, 2012, p. 381-382).

5.4.5 Durbin-Wu-Hausman Test

When working with panel data, it is standard to apply either a fixed- or random effects regression model. This is to control for unobserved heterogeneity. Because of this, the analysis in this study has to decide when to use a fixed effects or random effects model. This is done through a Durbin-Wu-Hausman test. According to Verbeek (2012, p. 385), a Durbin-Wu-Hausman test is for the null hypothesis that \( x_{it} \) and \( \alpha_i \) are uncorrelated. The
two estimators are compared: one is consistent under both the null and alternative hypothesis, and one that is consistent under the null hypothesis only. A significant difference between them indicates that the null hypothesis is unlikely to hold. In other words, if we run a Durbin-Wu-Hausman test we will get a p-value which indicates if a fixed effects or random effects model is most suitable to use. A Durbin-Wu-Hausman test will be performed for the regression models used in this study. The standard form of this test does not account for heteroskedasticity, which has to be taken into consideration when performing the test.

The main idea of the Durbin-Wu-Hausman test is that two estimators are compared. One is consistent under the null hypothesis only, and the other is consistent under both null and alternative hypothesis. A significant difference between the two estimators is a sign that the null hypothesis is not likely to be upheld. The test shows if the random effects and fixed effects estimators are significantly different (Verbeek, 2012, p. 385-386).

5.4.6 Breusch-Pagan / Cook-Weisberg Test for Heteroskedasticity

The relevance of this discussion stems from the fact that one of the assumptions in simple linear regressions is homoscedasticity. Basically, this says that the variance of the error terms should be constant. Ordinary Least Squares (OLS) is a method to estimate a linear regression model; it is one of the most basic concepts when learning about regression analysis. As long as the variance of the error terms is constant, which would mean that the homoscedasticity assumption is fulfilled, then there is no issue with using regression based on OLS. Brooks (2014, p. 180-187) states that if the data on the other hand is showing heteroscedasticity, then the standard errors need to be adjusted. The GARCH-family models use a process called maximum likelihood instead of OLS. This method finds the most likely values of the parameters based on the actual data.

Brooks (2014, p. 415-416) states that financial data generally has a number of tendencies that linear models fail to handle properly and thus speak for the use of ARCH/GARCH type models; First, leptokurtosis. That means that the distribution has fatter tails and a peakedness around the mean, higher than a normal distribution. Second, volatility is clustering, periods with high returns are expected to follow high returns, and vice versa. Third, linear models fail to account for the fact that volatility tends to increase more with large negative returns than a same size positive return (Brooks, 2014, p. 415-416).

Deciding which test to use for detecting heteroskedasticity in the sample is determined by how explicit we want to be about the form of the heteroskedasticity. As a broad rule, the more explicit we are, the more powerful the test will be; in other words, the more likely it will be to correctly reject the null hypothesis. That being said, if the heteroskedasticity takes a different form than what the test is able to handle, then heteroskedasticity might not be detected at all. The Breusch-Pagan test is a Lagrange multiplier test for heteroskedasticity. The main characteristics of Lagrange multiplier tests is that the model does not need to be specified under the alternative hypothesis and that the test is often simply a computation of the r-square value in an auxiliary function (Verbeek, 2012, p. 106). Tests such as these are designed to test for the null hypothesis of homoscedasticity against the alternative hypothesis of different types of heteroskedasticity (Verbeek, 2012, p. 105). If the null hypothesis of homoscedasticity is rejected, one option is to estimate a more general model which allows for heteroskedasticity (Verbeek, 2012, p. 198).
5.4.7 Collinearity and Multicollinearity

If correlation between explanatory variables is too high, this leads to multicollinearity problems. From this conclusions can be drawn about which variables are useful and which are not. By useful is meant to what degree a variable contributes to the regression model. If two variables are highly correlated, one of them should be removed, if the variables that are highly correlated are intended to measure the same thing. There is no particular limit to look for when deciding if there is too high correlation between variables (Studenmund, 2014, p. 272). Keeping both in the regression model could cause numerical instabilities, which would be a problem for the output of the analysis. Two variables being highly correlated is called collinearity, while multicollinearity is when a variable is highly correlated with two or more variables (Moore et al, 2011, p. 610). A perfect correlation is expressed by values of 1 or -1 in the correlation matrix, and 0 if the variables are completely independent of each other. Some statistical software will calculate a Variance Inflation Factor (VIF) value for each explanatory variable in a model. If a value above 10 is reached in the VIF measure, it is a strong indication that multicollinearity problems may exist between the variables in the model (Moore et al, 2011, p. 610).
6. Results

In this chapter, the research questions formulated in the Introduction will be answered. The presentation of the results will start with the presentation of the descriptive statistics and information about the distribution of the ESG Score variable. The base results will then be presented. Then a number of statistical tests will be presented. This leads the reader to the results that are derived from the chosen regression model and the answering of the research questions.

6.1 Descriptive Statistics

The descriptive statistics are presented in Table 5, with the purpose of bringing a general overview of the data used in the statistical analysis of this thesis. The GARCH (1,1) estimation of volatility has a higher absolute max value than the Realized Volatility, with a higher standard deviation. The GARCH (1,1) method of volatility measure has a minimum value of 6.90e-06, while the Realized Volatility has a minimum value of 0. The ESG Score has a distribution between 98.02 and 8.50, with a mean of 60.1343 and a standard deviation of 17.3652. The Volume variable fluctuates between a maximum value of 256 million traded shares and a minimum value of 0. The mean is 1.449 million and a standard deviation is 3.086 million. Market Capitalization varies between 727 619 million and 71.35 million, with a mean of 29 035.9 million and a standard deviation of 50 170.1 million. Total Debt to Enterprise Value fluctuates between a maximum value of 13.8117 and a minimum value of 0. The mean is 0.252621 and the standard deviation is 0.295759. The variable P/E Including Extraordinary Items has been altered in order to get rid of the risk that outliers may affect the result. P/E ratios above 100 and below -100 have been deleted, thus the variability is between 100 and -100. The mean P/E ratio is 17.5636 and the standard deviation is 17.7929. The S&P return has varied between a positive return of 6.83% and a minimum of 6.89%. The Realized Volatility of the S&P 500 is between 0.004755 and 0.

Table 5. Descriptive Statistics of Included Variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>MAX</th>
<th>MIN</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realized Volatility</td>
<td>1.34302</td>
<td>0.00000</td>
<td>0.000409</td>
<td>0.002564</td>
</tr>
<tr>
<td>GARCH (1,1) Volatility</td>
<td>3.515206</td>
<td>6.90e-06</td>
<td>0.0004134</td>
<td>0.0038054</td>
</tr>
<tr>
<td>ESG Score</td>
<td>98.0200</td>
<td>8.5000</td>
<td>60.1343</td>
<td>17.3652</td>
</tr>
<tr>
<td>Volume</td>
<td>256 000 000</td>
<td>0.00000</td>
<td>1 449 240</td>
<td>3 086 082</td>
</tr>
<tr>
<td>Market Capitalization (in Millions)</td>
<td>727 619</td>
<td>71.3500</td>
<td>29 035.9</td>
<td>50 170.1</td>
</tr>
<tr>
<td>Total Debt to Enterprise Value</td>
<td>13.8117</td>
<td>0.00000</td>
<td>0.252621</td>
<td>0.295759</td>
</tr>
<tr>
<td>P/E Including Extraordinary Items</td>
<td>99.9900</td>
<td>-99.9900</td>
<td>17.5636</td>
<td>17.7929</td>
</tr>
</tbody>
</table>
6.1.1 The ESG Score

The ESG Score variable used in this study has a wide range of values. As explained in the Method section, the possible interval for this score is 0 to 100, and can be seen as a percentage. A score of 0 is the worst score and a score of 100 is the best score. As mentioned, in Figure 4, descriptive values about ESG score can be observed. The highest ESG score registered is 98, the lowest is 8,5, and with a mean of 60. The standard deviation is 17.37. The skewness indicates a slightly longer tail of scores to the left side of the mean, while kurtosis shows a fatter tail to the right side of the mean. This can be seen in Figure 4, but is not is relevant for explaining how valid the distribution of ESG scores is. This is because the method behind the generation of ESG scores is what matters; if the method is considered valid, and is followed, then the exact distribution of scores is not of importance, other than having a distribution where there are both high and low ESG scores, allowing us to compare how volatility behaves when ESG score is higher or lower.

<table>
<thead>
<tr>
<th>S&amp;P 500 Return</th>
<th>0.068366</th>
<th>-0.068958</th>
<th>0.000430</th>
<th>0.010899</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500 Realized Volatility</td>
<td>0.004755</td>
<td>0.00000</td>
<td>0.000119</td>
<td>0.000305</td>
</tr>
</tbody>
</table>

Figure 4. ESG Score Distribution

In order to get an idea of how the ESG Score fluctuates over time within companies that generally perform well in ESG Score, Table 6 has been made. Companies that perform well as measured by the ESG Score, generally have a reasonably stable ESG Score over time. As can be seen in Table 6, the distributions for the companies are around 90 for
each year in the study. This is in general something that can be seen for companies that are performing well, the good performance is likely to persist and remain relatively stable over time. Of course, there are years where even well performing companies slip up, but as can be seen for PepsiCo Inc for example, the companies that have historically performed well seem to rebound from slumps in ESG Scores relatively fast.

Table 6. Example of High Scoring ESG Companies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3M ESG Score</td>
<td>86.50</td>
<td>90.32</td>
<td>91.19</td>
<td>91.22</td>
<td>90.56</td>
<td>90.76</td>
<td>85.97</td>
<td>90.60</td>
</tr>
<tr>
<td>Microsoft Corp</td>
<td>90.71</td>
<td>92.17</td>
<td>91.08</td>
<td>86.91</td>
<td>92.56</td>
<td>91.29</td>
<td>89.44</td>
<td>93.34</td>
</tr>
<tr>
<td>PepsiCo Inc ESG</td>
<td>88.12</td>
<td>90.01</td>
<td>92.27</td>
<td>82.91</td>
<td>83.23</td>
<td>83.07</td>
<td>79.04</td>
<td>87.86</td>
</tr>
<tr>
<td>Praxair Inc ESG</td>
<td>83.45</td>
<td>86.26</td>
<td>88.90</td>
<td>88.16</td>
<td>84.82</td>
<td>89.46</td>
<td>89.25</td>
<td>90.73</td>
</tr>
<tr>
<td>Texas Instruments</td>
<td>89.15</td>
<td>92.81</td>
<td>95.14</td>
<td>90.85</td>
<td>91.33</td>
<td>91.91</td>
<td>91.80</td>
<td>93.08</td>
</tr>
</tbody>
</table>

The average company in the sample, has an ESG Score of around 60. In order to get an idea of how company scores around 60 fluctuate Table 7 has been developed. As can be seen in Table 7, also for companies with ESG Scores around 60, the score is relatively stable over time.

Table 7. Example of Average ESG Scoring Companies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rockwell Automation ESG Score</td>
<td>70.93</td>
<td>70.44</td>
<td>63.64</td>
<td>65.46</td>
<td>66.51</td>
<td>66.04</td>
<td>70.39</td>
<td>66.94</td>
</tr>
<tr>
<td>Stryker Corp ESG Score</td>
<td>63.40</td>
<td>68.50</td>
<td>62.30</td>
<td>59.32</td>
<td>53.34</td>
<td>46.43</td>
<td>53.39</td>
<td>61.97</td>
</tr>
<tr>
<td>Molson Coors Breweries Co ESG</td>
<td>64.84</td>
<td>68.25</td>
<td>67.82</td>
<td>67.93</td>
<td>62.25</td>
<td>62.09</td>
<td>57.11</td>
<td>71.53</td>
</tr>
<tr>
<td>Qualcomm Inc ESG Score</td>
<td>65.33</td>
<td>69.62</td>
<td>70.84</td>
<td>81.61</td>
<td>79.10</td>
<td>66.54</td>
<td>71.27</td>
<td>70.54</td>
</tr>
<tr>
<td>Progressive Corp ESG Score</td>
<td>58.66</td>
<td>62.53</td>
<td>63.61</td>
<td>68.61</td>
<td>60.71</td>
<td>59.90</td>
<td>64.39</td>
<td>67.39</td>
</tr>
</tbody>
</table>
To get an idea of how the average ESG Scores of the entire sampling frame (504 companies) has developed over time, Table 8 was developed. To summarize, ESG Scores are relatively stable over time within companies. Across the sampling frame the average ESG Score has been steadily increasing from 56.22 in 2009 to 65.88 in 2016. Simultaneously, the number of companies with observed ESG Scores has also increased, from 420 in 2009 to 498 in 2016.

Table 8. Average ESG Score Development Over Time.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG Score</td>
<td>56.22</td>
<td>57.63</td>
<td>58.31</td>
<td>58.25</td>
<td>58.97</td>
<td>59.05</td>
<td>62.60</td>
<td>65.88</td>
</tr>
<tr>
<td>Observations</td>
<td>420</td>
<td>438</td>
<td>454</td>
<td>453</td>
<td>460</td>
<td>468</td>
<td>470</td>
<td>498</td>
</tr>
</tbody>
</table>

6.1.2 Correlation Matrix

Table 9 exhibits a correlation matrix of the independent variables used in the study. The highest correlation, 0.4244, is between the variables Trading Volume and Market Cap. ESG Score and Market Cap are correlated, at a value of 0.3468. The least correlated variables are S&P 500 Return and Total Debt to Total Enterprise Value, which at -0.0012 correlation are very close to independent of each other. As no values are near perfect correlation, we can conclude that there are no cases of high collinearity, hence multicollinearity problems are deemed unlikely. Thus the variables used in this study are not overlapping in their contribution to the model.

Table 9. Correlation Matrix.

<table>
<thead>
<tr>
<th>Variable</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ESG Score</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Volume</td>
<td>0.2153</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. Market Capitalization</td>
<td>0.3468</td>
<td>0.4244</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4. Total Debt to Enterprise Value</td>
<td>0.0524</td>
<td>0.0794</td>
<td>-0.0180</td>
<td>1.0000</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. P/E Including Extraordinary Items</td>
<td>-0.0840</td>
<td>-0.0952</td>
<td>0.0057</td>
<td>-0.0911</td>
<td>1.0000</td>
<td>-</td>
</tr>
<tr>
<td>6. S&amp;P 500 Return</td>
<td>-0.0021</td>
<td>-0.0158</td>
<td>0.0019</td>
<td>-0.0012</td>
<td>0.0047</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
6.2 Preliminary Regression Results

In Table 10, the results of the first set of regressions that were run can be seen. The R-square statistic is 0.0797 i.e. 7.97% for the random effects model and 0.0686 i.e. 6.86% for the fixed effects regression with Realized Volatility as the dependant variable. This means that around 7% of the Realized Volatility can be described by the model specified. The R-square for the GARCH (1,1) estimate of volatility is significantly lower, being 0.0039 and 0.0037 depending on the model. This means that this model is only capable of explain around 0.4% of the volatility, as estimated by GARCH (1,1). Here both random and fixed effects panel regression models are run for both forms of volatility used in the study. The most interesting variable, the ESG Score, is significantly negatively related to both the Realized Volatility and the GARCH (1,1) estimate of volatility. The size of the effect is around $4e^{-06}$, regardless of the regression model chosen at this stage. The size of the standard error varies slightly more than the average effect does. Other than the ESG Score, all of the control variables used in the study were statistically significant in the regressions presented in Table 10. Volume is positively related to Volatility. Market Capitalization is negatively related to Volatility, except for the regression run on Realized Volatility with fixed effects. Total Debt to Enterprise Value is significantly positively related to Volatility, regardless of model and estimation method of Volatility. P/E Including Extraordinary Items is negatively related to Volatility, regardless of model and estimation method. The S&P 500 Return and S&P 500 Realized Volatility are both positively related to Volatility. The number of companies in the regressions is the sample size of 481.

Table 10. Regression Results A.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Realized Volatility: Random Effects</th>
<th>Realized Volatility: Fixed Effects</th>
<th>GARCH (1,1) Volatility: Random Effects</th>
<th>GARCH (1,1) Volatility: Fixed Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG Score</td>
<td>-4.17e-06*** (2.80e-07)</td>
<td>-3.90e-06*** (3.02e-07)</td>
<td>-3.84e-06*** (4.65e-07)</td>
<td>-4.82e-06*** (5.88e-07)</td>
</tr>
<tr>
<td>Volume</td>
<td>1.81e-10*** (1.21e-12)</td>
<td>1.95e-10*** (1.28e-12)</td>
<td>7.40e-11*** (2.17e-12)</td>
<td>7.41e-11*** (2.47e-12)</td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>-5.66e-10*** (1.23e-10)</td>
<td>1.77e-09*** (1.48e-10)</td>
<td>-3.04e-09*** (1.83e-10)</td>
<td>-3.66e-09*** (2.87e-10)</td>
</tr>
<tr>
<td>Total Debt to Enterprise Value</td>
<td>0.0000832*** (0.0000108)</td>
<td>0.0000863*** (0.0000113)</td>
<td>0.0001011*** (0.0000193)</td>
<td>0.0000954*** (0.0000219)</td>
</tr>
<tr>
<td>P/E Including Extraordinary Items</td>
<td>-2.42e-06*** (1.48e-07)</td>
<td>-2.63e-06*** (1.50e-07)</td>
<td>-3.47e-06*** (2.80e-07)</td>
<td>-3.37e-06*** (2.92e-07)</td>
</tr>
</tbody>
</table>
Extraordinary Items

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P 500 Return</th>
<th>S&amp;P 500 Realized Volatility</th>
<th>Nr. Companies</th>
<th>Overall R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0012426***</td>
<td>0.0012913***</td>
<td>481</td>
<td>0.0797</td>
</tr>
<tr>
<td></td>
<td>(0.0001996)</td>
<td>(0.0001994)</td>
<td></td>
<td>0.0686</td>
</tr>
<tr>
<td></td>
<td>0.0019812***</td>
<td>0.0019689***</td>
<td></td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td>(0.0003884)</td>
<td>(0.0003884)</td>
<td></td>
<td>0.0037</td>
</tr>
<tr>
<td></td>
<td>1.693677***</td>
<td>1.695913***</td>
<td>481</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0072926)</td>
<td>(0.0072939)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.3648678***</td>
<td>0.3596055***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014171)</td>
<td>(0.0142033)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = 10% significance level ** = 5% significance level *** = 1% significance level.

6.2.1 Breusch-Pagan / Cook Weisberg Test for Heteroskedasticity

As discussed in section 5.4.6, failing to test for heteroskedasticity may lead to misleading standard errors in the regression model. Breusch-Pagan / Cook Weisberg tests have been performed to detect the presence of heteroskedasticity in the two models used; one model for each dependent variable. By using the command “hettest” in Stata 15, the test will be carried out on the linear regression which has just been computed in the program. The hypothesis used for this test is:

\[ H_A = \text{The error term has a non-constant variance.} \]

A 95% significance level is used, which means if the p-value (“Prob > chi2”) is above 0.05, then the null hypothesis (homoskedasticity) cannot be rejected, and if it is below 0.05 then the null hypothesis can be rejected. The results for the dependent variable GARCH (1,1) showed a Prob > chi2 of 0.0000 (Appendix 2). This means that the null hypothesis of constant variance can be rejected, and the GARCH (1,1) estimate of volatility has a non-constant variance. This means that there is indications of heteroskedasticity in the sample.

The same test was then run on the other dependent variable, Realized Volatility. The hypothesis for this test is:

\[ H_A = \text{The error term has a non-constant variance.} \]

The result for the dependent variable Realized Volatility, showed a Prob > chi2 of 0.0000 (Appendix 2). Similarly, as in the first instance, this means that the null hypothesis of constant variance can be rejected. There is presence of heteroscedasticity in the sample. This indicates that a robust regression model that handles heteroscedasticity better is in order for both dependent variables in this study.

6.2.2 Durbin-Wu-Hausman Test

As discussed in section 5.4.5, when using panel regression models, the decision whether to use fixed or random effects in the model is one issue (Greene, 2012, p. 419). The Hausman test is for the null hypothesis that \( \hat{\delta}_{it} \) (independent variable) and \( \alpha_i \) (intercept for each entity) are uncorrelated. A panel regression with random effects and robust
model, clustered by companyid, was performed, in order to be able to perform the Durbin-Wu-Hausman test. After that the Durbin-Wu-Hausman test was run using the “xtoverid” command, which is different from the “xthausman” which is used when there is no heteroskedasticity in the model. The hypothesis used for this test is:

\[ H_0 = \text{Random Effects is the preferred model} \]

The result of this test for the GARCH (1,1) estimate of volatility is a p-value of 0.2183 (Appendix 3). A 5% significance level is used; thus, the test indicates that the null hypothesis cannot be rejected. A Random effects regression model is the correct one to use for the dependent variable GARCH (1,1) Volatility estimate.

The same test was conducted for the Realized Volatility dependent variable. The hypothesis for this test is:

\[ H_0 = \text{Random Effects is the preferred model} \]

The result of this test for the Realized Volatility measure of volatility is a p-value of 0.0895. The same significance level of 5% is used here as well. This indicates that the null hypothesis cannot be rejected and a random effects model is the preferred model to use for this regression as well.

6.2.3 Testing for Multicollinearity

As explained in section 5.4.7, using a Variance Inflation Factor (VIF) test can reveal if there are multicollinearity problems in the specified model. A value above 10 is generally an indication that such problems may exits. A VIF test was carried out in order to test for multicollinearity between the independent variables, of which the results can be seen in Table 11. The results indicate that multicollinearity problems in the specified model are unlikely, as the highest VIF value is 1.34 for the variable Market Capitalization. The mean VIF value is 1.11.

Table 11. Variance Inflation Factor Test

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG Score</td>
<td>1.15</td>
<td>0.87</td>
</tr>
<tr>
<td>Volume</td>
<td>1.25</td>
<td>0.80</td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>1.34</td>
<td>0.75</td>
</tr>
<tr>
<td>Total Debt to Enterprise Value</td>
<td>1.01</td>
<td>0.99</td>
</tr>
<tr>
<td>P/E Including Extraordinary Items</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>S&amp;P 500 Return</td>
<td>1.01</td>
<td>0.99</td>
</tr>
<tr>
<td>S&amp;P 500 Realized Volatility</td>
<td>1.02</td>
<td>0.98</td>
</tr>
<tr>
<td>Mean VIF</td>
<td>1.11</td>
<td>-</td>
</tr>
</tbody>
</table>
6.2.4 Summary of the Conducted Tests

The regressions that are presented in Table 10 do not take the potential problems when using panel regression models into account. Subsequently, a number of tests were conducted in order to decide upon the most appropriate panel regression model to use. First the Breusch/Pagan–Cook-Weisberg test for heteroscedasticity revealed that there are indications of heteroscedasticity in both dependent variables, Realized Volatility and the GARCH (1,1) estimated volatility. This indicates that a robust covariance estimator needs to be used, in order to account for this finding, and receiving correct standard errors. Second, in order to decide on whether to use a random effects model or a fixed effects model, Durbin-Wu-Hausman tests were done. These tests indicated that a random effects model is the right choice in the case of both dependent variables. Finally, a VIF test was conducted to test for the risk of multicollinearity problems. This test revealed that the risk for this is small. These tests lead to the final regression models, the results of which can be seen in Table 12.

6.3 Final Regression Results

Table 12 shows the regression results obtained, using the models that have been deemed optimal for the data at hand.

Table 12. Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Realized Volatility: Random Effects Robust SE</th>
<th>GARCH (1,1) Volatility: Random Effects Robust SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESG Score</td>
<td>-4.17e-06*** (1.40e-06)</td>
<td>-3.86e-06*** (1.11e-06)</td>
</tr>
<tr>
<td>Volume</td>
<td>1.81e-10*** (2.68e-11)</td>
<td>7.41e-11*** (1.15e-11)</td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>-5.66e-10 (8.21e-10)</td>
<td>-3.03e-09*** (6.95e-10)</td>
</tr>
<tr>
<td>Total Debt to Enterprise Value</td>
<td>0.0000832 (0.0000993)</td>
<td>0.0001013 (0.000106)</td>
</tr>
<tr>
<td>P/E Including Extraordinary Items</td>
<td>-2.42e-06*** (5.74e-07)</td>
<td>-3.47e-06*** (5.61e-07)</td>
</tr>
<tr>
<td>S&amp;P 500 Return</td>
<td>0.0012426*** (0.0004291)</td>
<td>0.0019812*** (0.0002472)</td>
</tr>
<tr>
<td>S&amp;P 500 Realized Volatility</td>
<td>1.693677*** (0.0858389)</td>
<td>0.3648673*** (0.0273464)</td>
</tr>
<tr>
<td>Number of Companies</td>
<td>481</td>
<td>481</td>
</tr>
<tr>
<td>Overall $R^2$</td>
<td>0.0797</td>
<td>0.0039</td>
</tr>
</tbody>
</table>

*= 10% significance level **= 5% significance level ***= 1% significance level.

Table 12 displays the regression models that have been decided upon after appropriate statistical testing. The effect of ESG Scores on Realized Volatility is studied using a random effect, robust standard error panel regression. This model uses a robust covariance estimate to estimate the standard error of parameter estimates. The effect of
ESG Scores on the GARCH (1,1) Volatility is studied using the same random effects, robust standard error panel regression model.

Research Question 1 was introduced in the Introduction and was the following:

*Does a high ESG-score lead to lower share price volatility, as measured by daily GARCH (1,1) estimates of volatility, ceteris paribus?*

This research question can now be answered. The regression results indicate that ESG Score is statistically significantly negatively related to the GARCH (1,1) Volatility estimate. The relevant significance level used in the interpretation of the results is 5%, however, the results of the regression reveal that the significance level of the ESG Score is 1%, in both methods of volatility estimation. The coefficient is -3.86e-06 and the adjusted robust standard error is 1.11e-06. This indicates that on average, a firm that has an ESG Score of 90, compared to a firm with an ESG Score of 20, has a 0.0002702 lower volatility per day.

Research Question 2 was the following:

*Does a high ESG-score lead to lower share price volatility, as measured by simple realized daily volatility, ceteris paribus?*

This research question can also now be answered. The result of the regression indicates that the ESG Score is statistically significantly negatively related to Realized Volatility. The coefficient is -4.17e-06 and the adjusted robust standard error is 1.40e-06. What this means, is that on average each one step increase in the variable ESG Score, leads to a firm having 0.00000417 lower volatility per day. This implies that a firm with an ESG Score of 90, compared to a firm with an ESG Score of 20, has a 0.0002919 lower volatility per day, again, on average.

The regressions also revealed some other interesting results. The R-square is a measure which reveals how much of the variance in Y (the dependent variable) that can be explained by the regressors (the independent and explanatory variables used in the model) (Stock & Watson, 2015, p. 242). The R-square value for the Realized Volatility model, is 0.0797. This indicates that 7.97% of the variable Realized Volatility can be explained by the model specified. For the GARCH (1,1) Volatility the corresponding R-square value is 0.0039. This indicates that only 0.39% of the volatility of the variable GARCH (1,1) Volatility can be explained by the model that is specified here below. Volume is positively related to volatility, regardless of method of volatility estimation. The result is statistically significant. The variable Market Capitalization has lost its statistical significance under the Realized Volatility, but the variable is still statistically significantly negatively related to the GARCH (1,1) estimate of volatility. The variable Total Debt to Enterprise Value is no longer positively statistically significant to either volatility measure. P/E Including Extraordinary Items is negatively related to volatility. This result is statistically significant. The return on the S&P 500 is positively related to volatility, this result is also statistically significant. S&P 500 Realized Volatility is also positively related to both methods of volatility estimation. This result is also statistically significant.
7. Analysis

This final chapter will provide a discussion and an analysis about the results obtained from the regression model presented in the Results chapter. The results will be analyzed based on included theories, on the basis of previous research and from an investor perspective. The possibility of lurking reversed causality is discussed. The chapter is concluded with a discussion about ethical and societal implications of the research.

7.1 Discussion Surrounding the Research Results

The main result of this study, presented in the previous chapter indicates that high ESG performance lead to lower share price volatility between the years 2009 and 2016 on the S&P 500 firms studied. As can be seen in the regression output, the R-squared value obtained in the first model is 0.0039 overall and 0.0797 in the second model. This means that 0.39% of the volatility as estimated by GARCH (1,1) is described by the regression model. Furthermore, 7.97% of the simple realized volatility is described by the regression model used. In absolute terms this is considered low. However, consider that Roll (1988) found that explaining share price fluctuation is difficult in general. He looked at the effects of systemic economic events, other firms returns in the same industry and public firm-specific news announcement and found that the explanatory power was quite low. Roll, found a daily R-squared of 0.20. So, in comparison the model used in this study seems to be not too far off in terms of explanatory power. It is also important to consider that Shiller (1981) and LeRoy and Porter (1981) tried to use fundamentals to describe the fluctuation of share prices. DeLong et. al. (1990) and Campbell and Kyle (1993) find that noise traders affect prices on the market. Given the nature of noise traders, being very unpredictable in their trading behaviour, makes explaining the causes of volatility a fleeting task. A low R-squared does not invalidate the model, rather it shows, in addition to previous research that explaining volatility would require a very large model, with many variables, something that was not possible in the scope and timeframe of this study.

Furthermore, the difference in R-square can based on our models likely be described by the estimation method of volatility. Since all of the other variables are the same in the model, and the regression model is the same for both volatility estimates, this seems like a reasonable conclusion. The simple realized volatility is based on daily logarithmic returns, this is the actual observed volatility at each point in the past. The GARCH (1,1) estimate of volatility is conditioned upon several different inputs. The resulting measure thus takes a larger variability into account, and this larger variability was proven to be difficult to explain with the model used in this study.

Moreover, the results show that higher volume, leads to increased volatility. Smaller firms, as measured by market capitalization have higher volatility. The measure total debt to enterprise value was positively related to volatility, however this result was not statistically significant. A lower P/E ratio including extraordinary items generally leads to higher share price volatility. A higher return on the S&P 500 is also associated with higher volatility. The results are robust to the two different methods of volatility estimation in this study.
Volume leads to higher volatility. This is in line with the evidence of noise traders that influence the market by trading (DeLong et al., 1990; Campbell & Kyle, 1993). This is also in line with Fama (1965) and French & Roll (1986). Smaller firms, as measured by market capitalization, have higher volatility. Perez-Quiros & Timmermann (2000) found that there is asymmetric risk between large and small companies. Smaller firms were found to have a higher level of risk, which is in line with this study. The debt measure included in this study was not found to be statistically significant. This may be due to the S&P 500 companies being of generally sound financial health. If this was the case it would mean that even a relatively high debt to enterprise value, in this study would not be high enough for the company to be in financial distress. Moreover, the years 2009-2016 were a period of economic expansion, this makes it even more unlikely that a relatively high debt to enterprise value in our study would be synonymous with companies being in financial distress. The reasoning for including a debt measure can thus still be seen as reasonable. Furthermore, the results indicate that a low P/E ratio leads to higher volatility. A low (negative) P/E is held by companies that are making losses. This seems reasonable. As stated in the Method chapter, companies making losses may be viewed with scepticism and uncertainty about future earnings prospects as well. The one result that seems perplexing is that higher return on the S&P 500 is accompanied by higher volatility, however it may have a simple explanation in this study. Since this study was done during a market expansion, that means that there is relatively less downside volatility overall during this time-period, leading to this being perfectly reasonable.

7.2 Analysis of the Results on the Basis of Included Theories

Combining previous research into the financial performance of companies with high CSR performance with the result of this study seems to render some interesting results and implications for the efficient markets hypothesis and the CAPM. Previous research, Wang & Sarkis (2017), Rodriguez-Fernandez (2016), Filbeck et. al. (2009), McGuire et. al. (1988), Guenster et. al. (2005), Schnietz & Epstein (2005), Hillman & Keim (2001) and Turban & Greening (1997) has generally found that high CSR performance is associated with a higher financial performance. This result is however not unequivocal and there are also studies pointing to the opposite (Brammer et.al., 2006). However, on balance, research seems to indicate that high CSR performance leads to higher financial performance. So, pairing that result with the result in this study would render the conclusion that high CSR performance would mean both higher return and lower risk. This does sound too good to be true, and most things that sound too good to be true, are just that. Is this too good to be true? First, it is important to remember that this study was made only to study the risk component, so nothing can be said about the return of the shares in our sample. Second, the joint hypothesis problem should also be considered. The anomalous returns found in previous studies, are calculated by establishing some sort of baseline for expected return, in order to determine what part of the return is an anomaly (the part that is above expected returns). This baseline is calculated using some sort of asset pricing model, often CAPM. As mentioned in the theory chapter, the anomalous returns may be there, but it may also only indicate that the CAPM, which is often used as a proxy in order to find out what returns are anomalous, is unable to capture some aspect of idiosyncratic risk that supposedly then is higher, due to the higher returns. Of course, both can also be true, i.e. that the anomalous returns are there and the CAPM as an estimation model for expected return is faulty, according to the joint hypothesis problem, there is no way of knowing which is true.
Modern Portfolio Theory offers an explanation into how investors should choose portfolios under risk. The results of this thesis indicate that high ESG performance leads to lower share price volatility, this has a number of implications for how investors should choose portfolios. Of course, the choice of an investor's optimal portfolio is largely based on the risk appetite of the investor in question. MPT argues that investors should choose portfolios in a way that minimizes the risk at any given expected return of portfolios. Alternatively, the risk level that one aspires for can be chosen and the accompanied return with the chosen level of risk is then given. The efficient frontier along with the capital market line, defines the theoretically optimal portfolio to choose by investors. The market risk cannot be altered by the choice of portfolio, since it is an inherent trait in taking part in the stock market. There is evidence provided by previous studies that idiosyncratic risk factors matter to investors. For example, Campbell et.al. (2001) find that idiosyncratic risk has been increasing, when compared to market risk, in recent years. This would in turn mean less explanatory power for traditional market models such as the CAPM and the MPT. Ang et.al. (2006) argue that volatility i.e. total risk should be a priced factor in the market, not only the market risk. Furthermore, Kumar and Lee (2006) find that, by checking for retail investor concentration that shares with high arbitrage costs, and high idiosyncratic risk, react much stronger to changes in retail investor sentiment than other shares. Indicating that idiosyncratic risk is priced. Since high ESG scores have an effect of lowering the idiosyncratic risk, this in turn would lower total risk, everything else equal. These three papers (Campbell et.al., 2001; Ang et.al., 2006; Kumar & Lee, 2006), in conjunction with the MPT and the findings of this paper render interesting results. The findings of Ang et.al. (2006), Campbell et.al. (2001), Kumar and Lee (2006) suggest that idiosyncratic risk matters to investors. High ESG scores have been shown to alter total risk downward in this paper, thus this would imply that the efficient frontier should be altered, assuming that idiosyncratic risk matters (since that part of total risk is the most likely to have been altered with high ESG performance), as the three papers presented argue.

So, considering the wide benefits found from high ESG/CSR performance that have been shown in this and other studies, why do not all companies focus more effort on the matter? Also, why do not all companies achieve high ratings in CSR and ESG measurement?

The measurement of ESG and CSR performance is still very much an open discussion and a optimal way to measure it has not yet been found. This can be seen in the fact that there is a wide number of different sources used in the studies presented in the literature review chapter. For example, the Dow Jones Sustainability Index, Kinder, Lydenburg and Domini database, Domini social index, Fortune 500 survey on corporate reputations, Ethical Investment Research Service (EIRIS) and the Bloomberg ESG database are all used by different researchers in a wide variety of studies. Lee and Faff (2009) and Peloza (2006) recognize the problem with measuring ESG/CSR aspects of businesses as well. In addition, intuitively it is reasonable to think that measuring these aspects might create problems. For example, measuring something like workplace satisfaction, can reasonably be measured by a survey to employees asking them about the issue, this might render in biases, as it is subjective to each individual situation. Another approach is to measure for example absentee days, or workplace injuries. This would amount in a numerically measurable output, but some of the finer nuances might be missed. This measurement problem might deter companies from engaging fully in such activities. The logic behind this is that in order for there to be a financial incentive to engage in CSR activities, companies are unsure of what exactly to focus on. Consider for example the ESG score,
it includes within it a vast amount of different measures. This study does indicate that achieving a high ESG score leads to lower share price volatility. But, what aspects within this ESG score is driving the result of lower volatility? Is it the environmental part? The social part? Or is it the governance part? And then, which part within these three aspects? Moreover, if there is no certainty that the output can be measured, how can the company know that they will receive credit for their work? Until there is a more efficient method of making ESG output more tangible, the discussion of its usefulness is likely to persist.

One framework that could partially explain why some companies do not achieve high ESG scores is also The Life Cycle Sustainability Assessment Methodology (LCSA). This approach is being used by a number of companies already. Starting from the bottom, companies willing to focus on CSR might want to have an approach like the one prescribed by LCA and SLCA. This can amount itself in focusing on the entire life-cycle of products or production processes. This method is argued to be in use at Dow Chemical, as described by their CEO in a HBR Ideacast interview (Harward Business review, 2017).

Focusing on environmental aspects can give a solid foundation for achieving business objectives, and developing a competitive advantage through cost efficiency. SLCA is concerned with aspects such as human rights, workplace satisfaction. By focusing on aspects of LCA and SLCA, companies might also be able to get ISO certified. This could serve as a tangible proof of the work being done. Of course, the sole focus on environmental aspects is not in the best interest of profit seeking businesses. That means that combining the aspects of LCA and SLCA with LCC, life-cycle cost analysis could be beneficial. By managing the aspects of LCA, SLCA and LCC well, it could be feasible to argue that companies that are able to do this, might be able to achieve high ESG scores.

This makes sense when considering that many of the aspects that are included in the LCSA are the same that many CSR and ESG scoring systems look at. However, when looking at Figure 2, showing the main differences between LCA and LCC, it becomes clear that some of the aspects seem difficult to manage simultaneously. So, it might be feasible to think that some companies are able to manage these tradeoffs better than others. Thus such companies may achieve higher ratings in for example ESG scores, and through this be able to reap the rewards of the lower volatility that has been found to be the case in this study. So, in order to achieve the full benefits of investing in CSR aspects, the key may be to manage the tradeoffs between the short termism of financial profit pursuit with the long term outlook of focusing on CSR aspects.

7.3 Results Alongside Previous Research

The results of this thesis can be seen as being in line with the findings of Harjoto et.al. (2017), Mishra and Modi (2013), Jo & Na (2012), Lee & Faff (2009) and Luo & Bhattacharya (2009), Spicer (1978). However, the studies mentioned investigated CSR effect on idiosyncratic risk, apart from Harjoto et.al. (2017), the result is the same, in the sense that good CSR performance lowers risk.

The chosen topic overlaps four general areas of research, and each of these have been discussed throughout this thesis. First, research into what aspects cause volatility (Shiller (1981); LeRoy and Porter (1981); Roll (1988); DeLong et.al (1990); Campbell & Kyle (1991)) and a discussion about idiosyncratic risk (Campbell et.al. (2001); Wei and Zhang (2006); Kumar and Lee (2006); Ang.et.al. (2006)) was included in the Theoretical Framework chapter. Second, research investigating the relation between CSR performance and risk, such as Harjoto et.al. (2017), Nguyen et.al. (2015), Mishra & Modi
Third is the largest body of research, investigating the relation between CSR and corporate financial performance this includes studies by for example Wang & Sarkis (2017), Rodriguez-Fernandez (2016), Filbeck et.al. (2009) and Brammer et.al.(2006). The risk-return tradeoff along with the research in that area was discussed in conjunction with efficient markets and CAPM above, as they are highly interconnected issues.

The fourth body of research is concerned about the CSR-cost of capital relationship. In this area studies by Suto & Takehara (2017), Chava (2014), El Ghoul et.al. (2011) and Sharfman & Fernando (2008) find that higher CSR performance lowers cost of capital. Cost of capital is generally a weighted average of the company cost of debt and cost of equity, as measured in the studies presented. The concept was developed by Modigliani and Miller, in their 1958 paper *The Cost of Capital, Corporation Finance and the Theory of Investment*. In the paper, Modigliani and Miller (1958) present three theorems, first that capital structure is irrelevant. Second, the cost of equity capital is increasing with the level of debt, the additional debt makes the equity more risky. Theorem three identifies the weighted average cost of capital as the hurdle rate for companies and adds that the type of funding used is irrelevant under the assumptions put forth for the three theorems. The findings of Suto & Takehara (2017), Chava (2014), El Ghoul et.al. (2011) and Sharfman & Fernando (2008) thus indicate that, similarly to this study, that high performance in CSR lowers risk. The difference is that cost of capital is on the firm level, so risk is then lowered within the firm. The determinant of the weighted average cost of capital is in the end the debt and equity markets, and their view of the riskiness of the firm. The measure that has been employed in this thesis measures the view of the investors of the effect of high CSR performance on share price volatility. Furthermore, since the effect on CSR should plausibly first be visible within the firms conducting CSR, before showing up on the stock market, the research within CSR-cost of capital can be viewed as reinforcing the result of this study. In other words, the research pointing to the possible fact that CSR lowers firm cost of capital also points to the plausible scenario that CSR does fill a real function within companies, by lowering their cost of capital.

As mentioned, the aspects that cause volatility are, based on previous research very difficult to pinpoint with accuracy. Fundamentals, maybe unsurprisingly, have been shown to affect prices of shares and their fluctuation. Shiller, 1981 showed that prices tend to overreact to dividend changes. In this study, at least P/E ratio and volume has been shown to affect volatility, regardless of volatility estimation method. Part of the reason that volatility is so difficult to pinpoint may be that investors have very different models to predict future stock prices, and true intrinsic values of shares. Based on this thesis, as well as Harjoto et.al. (2017), Nguyen et.al. (2015), Mishra and Modi (2013), Jo & Na (2012), Lee & Faff (2009) and Luo & Bhattacharya (2009) ESG/CSR factors should be included in these fundamentals that affect volatility.

ESG performance is a firm specific measure, and as such it seems to lower idiosyncratic risk. Good management, as defined by a high ESG score, does seem to lead to benefits that lower overall volatility. As mentioned earlier in the analysis, the market risk is unlikely to change, and this would then suggest that idiosyncratic risk is altered by high ESG performance. Wei & Zhang (2006), Kumar & Lee (2006) and Ang et.al. (2006) suggest that this factor has been increasing in importance and does, contrary to traditional theory, affect investor behaviour.
7.4 The Results from an Investor Perspective

This issue was touched upon already in chapter 7.2 but will be discussed more in depth here. The MPT states that the rational investor is risk averse, this means that they should strive to minimize the risk at any given level of return. The characteristics of investors portfolios look different in reality, however. Private investors trade too much and underperform (Barber & Odean, 2000), prefer to own shares familiar to them (home bias) (French & Poterba, 1991; Grinblatt & Keloharju, 2001; Kilka & Weber, 2000) and hold portfolios that include fewer stocks than is optimal by traditional theory (Goetzmann & Kumar, 2008; Polkovanichenko, 2005). Keep in mind that traditional theory such as MPT and CAPM suggest that idiosyncratic risk does not matter. One of the assumptions of both these theories is that investors are rational risk minimizers and diversify away the firm specific risk. If they then in reality do not diversify as much as is proposed and are biased to choosing domestic stock leading to underdiversification, it means that idiosyncratic risk is reasonably sure to prevail in their portfolios. This, in turn would mean that being able to lower idiosyncratic risk by owning shares in firms that have high ESG scores would be highly attractive. Owning shares in companies that have high ESG performance would then lower the overall risk of the portfolio, everything else equal. But, it is by no means a solution that magically leads to full diversification, if investors still diversify too little and, for example, own too large a share of domestic stocks, there is only so much that owning shares of high ESG performing companies can do on its own.

7.5 Direction of Causality

When conducting research on ESG/CSR and its effect in relation to financial performance or any other aspect of a business, the direction of causality is worth discussing. When seeing a statistically significant relation between a dependable and independent variable, this says nothing about causality. In the case of this study, it would be to pose the question if the reverse was true about the result derived. Would it be possible for a company to improve ESG performance by decreasing share price volatility? While this reversed order of dependent and independent variable would be possible to use in a study, the problem of which variable is responsible for change in the other variable is still present. This comes down to logic, as a regression analysis cannot give an answer about the direction of causality. In one type of hypothesized experiment, it might be easy to conclude that one variable is responsible for the change in the other variable, and the reverse cannot be true. For example, excessive abuse of alcohol will have an effect on which condition your liver is in, but it can be ruled out that a liver of poor health does not cause its owner to start consuming alcohol excessively. This is a direction of causality which can easily be proved. In this study, the result indicates that high performance in ESG aspects lead to lower share price volatility. This result can also plausibly be because firms that have high ESG scores perform well economically and thus have more available resources to spend on CSR activity, leading them to have high ESG scores. Carroll (1991) lays forth the The Pyramid of Corporate Social Responsibility model, which is a hierarchy of the components of CSR. The components are, in order of required for the next to be worked on: economic, legal, ethical, philanthropic. While all of them are highly desirable for a company, it is in general required to fulfill one component to be able to focus on the one above it. Carroll (1991) especially emphasizes that a business has the be profitable, which is the most basic step, to be able to engage in the other components of CSR. Assuming
that companies follow this logic when deciding upon CSR policy, and linking this with the result from this study, more profitable companies should on average have lower share price volatility, because they are able to focus more on ESG related matters. The reverse is also true, companies performing poorly in economic terms are unable to free the resources to engage in CSR related activity. This is one alternative way of explaining the direction of causality for the result in this study. It also is one possible explanation for the dispersion in ESG scores across firms.

7.6 Societal and Ethical Implications of the Research

7.6.1 Ethical Implications

The ethical considerations were presented in the Scientific Method chapter. This study is based on secondary data, thus the authors were not present during the data collection process. This means that there is no way of knowing how the data was initially collected and to what degree participants were informed, and shielded from harm. The data that is analyzed by Thomson Reuters is however public and that means that the data has been submitted by the companies either voluntarily or due to legal requirements. This makes it unlikely that any infringements have been committed, when it comes to privacy or harm to participants. The authors of this study have been truthful and done their best in trying to avoid misleading, or misrepresenting the results, contents of the data or the data management process. The research has rendered in increased understanding about the issues studied, meaning a reciprocal benefit for society, as well as increased knowledge for the authors. The authors have presented their background in the subject. This is done in the interest of honesty and transparency. The lack of professional experience within the field might be a downside. On the other hand, that might mean being able to approach the issues without any preconceived notions about the process, subject or results and how they relate.

7.6.2 Societal Implications

The research indicates that a high ESG Score lowers volatility. There is a large and growing interest in sustainable investing and ESG is the second largest aspect within the discipline according to GSIA. This study reaffirms that the overall increasing stream of investor money into sustainable investment projects seems to be the right way to go. It also gives further arguments for governments and other entities with a supporting role in the economy to keep supporting and to increase the commitment to sustainable projects and grants for companies to conduct business in a way that is sustainable. The European Commission identifies CSR as important to society due to it offering values that can build a more cohesive society onto which a sustainable economic system can be built (European Commission, 2018). The EC further states that it has a strategic interest in enhancing market rewards for CSR management of businesses, to improve company disclosure in the area, to further integrate CSR into education, training and research. This piece of research adds to research in the CSR area naturally.

It was also shown in Table 8 that the amount of companies covered by the Thomson Reuters ESG Score has increased from 2009 to 2016. This can be due to a few different reasons. It might be a case that companies are reporting their CSR efforts more extensively, which would be in line with the strategy of the European Commission. Gamerschlag et.al. (2011) provides evidence suggesting that companies with a higher
exposure to US shareholders increasingly disclose CSR aspects. Furthermore, the authors argue that in order to lower political cost, companies have an incentive to disclose as much CSR aspects as possible. Lower political cost implies reducing the potential impact of additional regulation, taxes and other aspects that may put downward pressure on firm value. The opposite also applies, higher political cost, due to lack of reporting, might lead to increased obstacles for business (Gamerschlag et al., 2011, p. 257). This reasoning goes hand in hand with the objectives of the European Commission. It is of course also possible that Thomson Reuters have missed out on some data due to some reason. This explanation seems more unlikely, since the authors of this paper are unaware of any changes made to the methodology for the period of 2009 to 2016. By including the LCSA framework into this thesis, as the reader may recall, one of the main issues in the framework is the acknowledgement that economic interest and holistic processes, such as CSR need to be managed simultaneously. Furthermore, the framework highlighted another major issue, which is the short-term horizon for profit maximization and the long-term view needed to make CSR work. Porter and Van Der Linde (1995, p. 123) provide a tangible example of how this clash between short-term profit seeking and the long-term nature of CSR might present itself. The example is from the pulp- and paper industry, where the problem has been the bleaching of paper with chlorine, this process harms the environment. A solution was found, in form of better processes in cooking and washing of the pulp, the chlorine was eliminated by using oxygen, ozone or peroxide for the bleaching, this was however not problem-free still. In return the production costs could be lowered (can be seen as a long-term benefit), however, the chlorine free paper sold at a 25% premium compared to the regular chlorine-bleached paper. Reasonably, this may hurt short-term profits, as some of the demand might be lowered due to the higher price (Porter & Van Der Linde, 1995, p. 122-124). The European Commission strives, with their strategy to support businesses in these issues, to solve some of the conflicting objectives that arise when trying to manage CSR alongside a profitable business (European Commission, 2018).

The results of this study, that high ESG performance leads to lower volatility adds to factors that make sustainable investment attractive, both on an investor and firm-level. This might give asset management companies, banks and other financial entities added incentives to create funds that are managed sustainably. Companies may benefit from a more stable share price, making for example stock-option incentives less risky. The increasing flow detected by GSIA into sustainable investment by 25% between the years 2014 and 2016 suggests that this process is already well underway. Again, this paper further strengthens the notion that this direction of flows of funds seems to be correct. The objectives set forth by the EC and the increase of funds in sustainable investment can be seen as a self-reinforcing process. The more money that goes into sustainable investment, the more the investments and projects associated with it are likely to grow, as funding is one of the key determinants to the success of any project. By investing sustainably this self-reinforcing cycle can be sustained.
8. Concluding Chapter

In this chapter, the final thoughts about this study and its results will be presented. The generalizability and the limitations of the study will be discussed. Truth criteria will be evaluated, to provide insight into how reliability and validity is handled throughout the study.

8.1 Conclusion

The aim of this study was to answer the research questions that were proposed by the authors. The research questions were the following;

Research Question 1: Does a high ESG-score lead to lower share price volatility, as measured by daily GARCH (1,1) estimates of volatility, ceteris paribus?

Research Question 2: Does a high ESG-score lead to lower share price volatility, as measured by simple realized daily volatility, ceteris paribus?

These research questions were answered by running two separate robust, random error panel regression analyses, after making appropriate statistical tests to determine the models to be appropriate. To clarify, one regression for the dependent variable simple realized volatility and a separate regression for the GARCH (1,1) estimated volatility were done. It is evident in the results section that there is a statistically significant negative relation between both forms of estimated volatility and ESG-score.

The analysis of the results raises the question if the traditional asset pricing models, CAPM and the portfolio theory model, MPT are really capturing all risk factors that are relevant in determining the risk-return relationship. The results of this study suggest that good ESG performance is able to lower the total volatility of shares. As discussed, the market risk is unlikely to have changed, it is likely the idiosyncratic part of total risk that has been altered. This result coupled with the investor portfolio characteristics and the recent studies indicating that idiosyncratic risk is rising and matters to investors renders us as authors of this paper to believe that the CAPM and MPT do not capture all relevant risk factors.

8.2 Theoretical and Practical Contribution

The theoretical contribution of this thesis is at least the following: This study has added to the literature that studies the ESG/CSR relation to volatility i.e. total risk. There are surprisingly few studies in this area, so this study will increase the depth of research in the area. The ones that exist are largely based at least partially on surveys and are older, as well as focused mostly on idiosyncratic risk. This study has used a more appropriate form of CSR data, as explained in the critique on previous research. Previous studies use CSR measurement tools that have absolute bars of performance, this leaves a part of the sample out, as previously explained. This study thus offers a more complete measure of ESG performance. Furthermore, this study uses two different methods of volatility estimation, and can present results robust to both models of estimation. The result
presented, that high ESG performance lowers share price volatility adds to the discussion of whether or not total risk should be a priced risk factor on the market. The results of this study, accompanied by previous research add to the evidence pointing to that total risk should be priced.

The practical relevance of this study is the following: The study provides added knowledge into the portfolio selection process. The study indicates that risk can be lowered by holding companies that perform well in ESG aspects, assuming that idiosyncratic risk factors are priced in the market, as certain studies suggest. Even in the situation that idiosyncratic risk wouldn’t matter, given the portfolio characteristics of individual investors, not being fully diversified and some idiosyncratic risk still remaining in investors portfolios, the results of this indicates that the risk can be lowered by holding shares of high ESG performing companies.

The study provides a further incentive for companies to focus on CSR. The study indicates that risk can be lowered by performing well in CSR aspects. This indicates that companies that strive for stable stock market performance, may benefit from investing in CSR. Furthermore, some possible explanations and issues were highlighted that may help companies become aware of aspects that may inhibit performance in CSR.

The study further reaffirms that the flows of capital into sustainable investment seen in recent years, is appropriate and should be sustained, due to the numerous benefits that previous and this research suggest are there.

8.3 Limitations of this Study

As with any research in the area of business, this study does have some limitations. First, the choice of using daily volatility estimates instead of the more commonly used yearly or monthly estimates. While it might give a better estimate of volatility, due to a higher frequency of measurement, it also might hinder direct comparison to other studies that use yearly or monthly volatility estimates. This choice has been argued for in the delimitations. Second, the measurement problem described throughout this thesis is not necessarily solved by the Thomson Reuters ESG score. While the authors of this paper view it as a better and more appropriate source of data than what has previously been available, there is no guarantee that it is completely unbiased, or even better. Third, the data of this study is based on solely secondary sources, this hinders the authors to check for ethical factors in the collection of the data, and making sure that it is assembled in a fair manner. It also inhibits the authors from proofing the integrity of the data. Fourth, a few of the control variables included in the model were found to be insignificant, this may be due to the way they are measured and another measurement for debt than Debt to Total Enterprise Value might have yielded different results. This does, however, question the decision on including those specific variables in the study.

8.4 Suggestions for Future Research

Based on the experience gained while writing this paper, the authors, Jakobsson and Lundberg would like to suggest the following for future research:
The companies used in this study have all been centered on the United States. The study should be attempted with data centered on companies of a different geographical region. The authors of this study are aware that altering the geographical association for the sample might be hard to do, as data of comparable quality and quantity may be scarce in supply. It is of interest to see if the results of this study still are valid if using companies from a different part of the world; for example, a study using the same methods, but with data from European companies, would be of interest.

Employing a different basis for ESG evaluation, namely replacing the Thomson Reuters ESG-score with data collected using a different set of evaluation criteria / methods. Which is to say, there are different companies evaluating companies’ ESG performance, and it is of interest to see if there are significant differences in results when using another company's data for ESG performance. Due to time constraints, and an apparent lack of consensus about variables affecting share price volatility, this study did not include a large number of explanatory variables in the regression analysis. Future research should absolutely attempt to include more explanatory variables, as controlling for more factors affecting a company's share price volatility is desirable.

Performing this study with data from a different time span is of interest, as the period used in this study, 2009-2016, has been a period of market expansion. For example, it is possible to see different results if a financial crisis were to be included in the time period used in the study, or simply using a period of market expansion in years other than 2009-2016.

In section 7.2. the question of which aspect of the ESG score drives the result obtained in this thesis, future research could plausibly split the ESG score into its components, in order to gain insight into the open questions presented in chapter 7.2.
9. Truth Criteria

9.1 Reliability

Reliability is concerned with the precision and accuracy of the measurement and absence of differences in the results if the research were repeated (Collis & Hussey, 2014, p. 52). Reliability is of particular importance when conducting quantitative research, since a researcher using a quantitative method is likely concerned with the question of stability when doing measurements. For example, if a scale was constructed to measure the weight of stones, reliability would not be high if the test was conducted the same way again and yielded different results for the same stones. There would be serious doubts about the test, and thus it would be deemed an unreliable measure (Bryman & Bell, 2015, p. 49).

In the case of this study, one half of the data consists of publicly available information about companies, which will not be altered or corrected as time passes. This is for example daily closing price for a company share, or the trading volume, which is an objective measure. The other half of the data used is information collected by an agency, Thomson Reuters, to compile an aggregate annual score to measure ESG performance. This ESG-score data used will also not be altered with time. However there is no guarantee that Thomson Reuters will not alter the method by which they compile this score, opening the possibility for an altered composition of weights in the future scoring system. The problem of lack of consensus in evaluation methods of ESG performance has been touched upon in this thesis, which could affect the grading of the same company making ESG investments, depending on the agency grading these ESG investments. Which means that from a data standpoint, there could be differences in the result if another rating system for ESG performance has been used. The publicly available data for the companies of the S&P 500 does not have altering methods depending on source, and thus should not affect the results of a replicative study. The statistical methodology employed should not alter the results of a replicative study, unless mistakes are present in the coding of data. Altering the chosen statistical software for the analysis should not make a difference.

9.2 Validity

Validity is the extent to which a test measures what the researcher wants it to measure and the results reflect the phenomena under study (Collis & Hussey, 2014. p. 53). There are a few different aspects of validity which is of particular interest when conducting quantitative research:

Measurement Validity, or at times called Construct Validity, concerns if measuring something really represents the concept which it intends to be investigating. For this to hold, it is presupposed that the underlying methods of measurement are reliable (Bryman & Bell, 2015, p. 50). In this study this is exemplified by 1) the two measurement types for volatility: simple realized volatility, and the GARCH(1,1) estimated volatility. 2) the measurement of ESG performance, the Thomson Reuters ESG-score. 3) the control variables, such as trading volume and total debt to total enterprise value. All of these variables used in the study are considered to give good measure of what they intend to convey.
Internal Validity deals with the matter of causality in a study. If two variables are measured, and they are said to have a causal relationship, then we must be confident of this. Is the independent variable X really causing a change in the dependent variable Y, or is the explanation for the change in Y a completely different one, that the researcher has failed to account for? (Bryman & Bell, 2015, p. 50). In the case of this study, the absolute measure would be the coefficient for the variable in the statistical result, and the accompanying measures statistical significance for these variables. These will either with a 95% or a 99% confidence level show that variable X affects Y to a certain degree, or if not having an effect at all. Valid results were found and presented, but causality cannot be proven, since regression analysis can only prove a relation, not causality. The causality linked to these results was discussed in the Analysis part.

External Validity is concerned with the question of results being generalizable beyond the specific research context. External validity is one of the main reasons researchers are keen to generate representative samples (Bryman & Bell, 2015, p. 50 - 51). There is a certain amount of generalizability outside the sample in this study. See the next part, 9.3, Generalizability of the results, for a complete exploration of this topic.

9.3 Generalizability of the Results

In this section, the discussion will surround the issue if the result of this study can be generalized further than the S&P 500 index, which includes medium and large US companies. This study includes a total of 481 companies from the index. The total number of companies included in the S&P 500 index is 504. The S&P 500 is divided into sectors on the basis of what industry they do business in. GICS is as mentioned used for the division and includes 11 sectors. It is reasonable to think that these same sectors are represented on every major stock market index as well.

The ownership structure on the S&P 500 should be considered when pondering the generalizability of the results. This is because, as has been argued in the thesis individual and institutional investors behave differently. According to S&P Capital IQ (2014) the S&P 500, just as the rest of the US stock market is largely owned by institutional investors, 82% of outstanding shares on the S&P 500 was owned by institutional investors in 2014.

So, the diverse nature of the S&P 500, including multiple companies from every sector from the GICS makes it likely that this study is generalizable to other stock markets, provided that certain characteristics, such as company size, diversity of industries, and ownership structure as the S&P 500 exhibits.
List of References


Appendix

Appendix 1. List of Deleted Companies

<table>
<thead>
<tr>
<th>Company Name</th>
<th>GICS Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accenture Plc</td>
<td>Information Technology</td>
</tr>
<tr>
<td>Align Technology</td>
<td>Health Care</td>
</tr>
<tr>
<td>Brighthouse Financial Inc</td>
<td>Financials</td>
</tr>
<tr>
<td>Chesapeake Energy Corporation</td>
<td>Energy</td>
</tr>
<tr>
<td>Crown Castle Corporation</td>
<td>Real Estate</td>
</tr>
<tr>
<td>DowDuPont</td>
<td>Materials</td>
</tr>
<tr>
<td>Fortive Corporation</td>
<td>Industrials</td>
</tr>
<tr>
<td>Garmin Limited</td>
<td>Consumer Discretionary</td>
</tr>
<tr>
<td>Intercontinental Exchange Inc</td>
<td>Financials</td>
</tr>
<tr>
<td>IPG Photonics Corporation</td>
<td>Information Technology</td>
</tr>
<tr>
<td>Iron Mountain</td>
<td>Real Estate</td>
</tr>
<tr>
<td>JM Smucker Co</td>
<td>Consumer Staples</td>
</tr>
<tr>
<td>KraftHeinz Inc</td>
<td>Consumer Staples</td>
</tr>
<tr>
<td>Pentair Plc</td>
<td>Industrials</td>
</tr>
<tr>
<td>Perrigo Company Plc</td>
<td>Health Care</td>
</tr>
<tr>
<td>Qorvo Inc</td>
<td>Information Technology</td>
</tr>
<tr>
<td>Regeneron Pharmaceuticals Inc</td>
<td>Health Care</td>
</tr>
<tr>
<td>SVB Financial Group</td>
<td>Financials</td>
</tr>
<tr>
<td>Take-Two Interactive Software Inc</td>
<td>Consumer Discretionary</td>
</tr>
<tr>
<td>TechnipFMC</td>
<td>Energy</td>
</tr>
<tr>
<td>Under Armour (ticker UA.N)*</td>
<td>Consumer Discretionary</td>
</tr>
<tr>
<td>Vertex Pharmaceuticals</td>
<td>Health Care</td>
</tr>
<tr>
<td>WestRock Company</td>
<td>Materials</td>
</tr>
</tbody>
</table>

*Under Armour ticker UAA.N has been included in the study, while UA.N has been deleted.
Appendix 2. Breusch-Pagan / Cook-Weisberg Tests

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of GARCH (1,1) Volatility

\[
\begin{align*}
\chi^2(1) &= 153.41 \\
\text{Prob} > \chi^2 &= 0.0000
\end{align*}
\]

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: fitted values of Realized Volatility

\[
\begin{align*}
\chi^2(1) &= 9.59e+06 \\
\text{Prob} > \chi^2 &= 0.0000
\end{align*}
\]

Appendix 3. Durbin-Wu-Hausman Tests

For GARCH (1,1) Volatility dependent variable:

Test of overidentifying restrictions: fixed vs random effects
Cross-section time-series model: xtreg re robust cluster(companyid)
Sargan-Hansen statistic 8.281 Chi-sq(6) P-value = 0.2183

For Realized Volatility dependent variable:

Test of overidentifying restrictions: fixed vs. random effects
Cross-section time-series model: xtreg re robust cluster(companyid)
Sargan-Hansen statistic 10.965 Chi-sq(6) P-value = 0.0895

Appendix 4. ESG Score Weights

<table>
<thead>
<tr>
<th>Pillar</th>
<th>Category</th>
<th>Indicators in Rating</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental</td>
<td>Resource Use</td>
<td>10</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>Emissions</td>
<td>22</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>Innovation</td>
<td>20</td>
<td>11%</td>
</tr>
<tr>
<td>Social</td>
<td>Workforce</td>
<td>29</td>
<td>16%</td>
</tr>
<tr>
<td></td>
<td>Human Rights</td>
<td>8</td>
<td>4.50%</td>
</tr>
<tr>
<td></td>
<td>Community</td>
<td>14</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Product Responsibility</td>
<td>12</td>
<td>7%</td>
</tr>
<tr>
<td>Governance</td>
<td>Management</td>
<td>34</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>Shareholders</td>
<td>12</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>CSR Strategy</td>
<td>8</td>
<td>4.50%</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>178</td>
<td>100%</td>
</tr>
</tbody>
</table>