A study of the relationships between corporate social performance, financial performance, and idiosyncratic risks

Author: Mateusz Rosół

Supervisor: Cristiana Parisi

Copenhagen Business School

Copenhagen, Denmark

A master thesis submitted in partial fulfillment of the requirements for the degree of Master of Science (MSc) in Economics and Business Administration (Finance and Investments)

Number of pages: 80

Number of characters (including spaces): 176,533

Date of submission: January 15, 2018

Abstract

This thesis studies the relationships between corporation social performance (CSP), corporation finance performance (CFP), and firm risk. The relations between ESG measures, market returns, and idiosyncratic risks were analyzed using a panel set of European firms during the period 2001-2016.

An investigation was conducted to determine any distinction between the relationship shared by ESG measures and downside idiosyncratic risk in comparison to that between ESG measures and upside idiosyncratic. There does not appear to be a distinction in our data.

A mediation analysis was conducted to determine whether there was any evidence that the relationship between ESG measures and market returns is mediated by idiosyncratic risk. The results were inconclusive.

Table of contents

1 Introduction

The Global Sustainable Investment Alliance (GAIA) reported in 2016 that $22.89 trillion of assets are managed using responsible investment strategies, representing an increase of 25% since 2014. This translates to responsible investment accounting for 26% of all assets under management globally (Global Sustainable Investment Alliance 2016).

Despite this clear demand from investors, there appears to be little consensus regarding the definition of responsible investing, or the related socially responsible investing (SRI) and impact investing, and their relation to financial performance (Milken Institute 2012). The same appears to apply to the broader concepts of corporation social responsible (CSR) and sustainability (Wang, Dou, & Jia, 2016).

The aim of this thesis is to reveal some of the related dimensions contained within the broad concepts of corporation social responsibility (CSR), corporation financial performance (CSP), and risk, and examine the theoretical and empirical relationships between them.

1.2 Problem formulation

The following line of questioning motivates this study: "Does CSP affect the risk of a firm, and if so, does it equally affect downside risk and upside potential? Having established this, can these effects explain the relationship, if one so exists, between CSP and firm returns?"

Placed in context, this thesis aims to address one of the knowledge gaps in the CSR discourse revealed in the review conducted by Aguinis & Glavas (2012): an understanding of the underlying mechanisms linking CSR and its outcomes. They found that only 7% of studies sampled explored mediation effects, despite the considerable study of the link between CSR and organizational outcomes. They developed a multi-level and multidisciplinary theoretical framework that organizes CSR research into institutional, organizational, and individual levels of analysis and organizes variables of study into categories consisting of predictors of CSR, outcomes of CSR, mediators of CSR-outcomes relationship, and moderators of CSR-outcomes relationship.

This analysis is situated within their framework as a study at the organizational level of risk as a mediating variable on the relationship between CSR and the outcome of CFP.

However, prior to such an investigation, consideration should be given to the observation that CSR, CFP, and risk are each complex multidimensional constructs. Research has begun to untangle the differing relationships between the distinct but related dimensions enveloped by each broader construct (e.g., Oikonomou, Brooks, & Pavelin 2012; Sassen, Hinze, & Hardeck 2016; Bouslah, Kryzanowski, & M'Zali 2013). ESG measures, market returns, and idiosyncratic risk were the dimensions chosen to investigate within the broader constructs of CSR, CFP, and risk, respectively. Once defined, the downside and upside dimensions of risk are investigated.

1.2.1 Research question

Following from the problem formulation, the main research questions are follows. Each of which in turn consists of several sub-questions that must be consecutively addressed.

1. Is there a discernable distinction between the relationship shared by ESG measures and downside idiosyncratic risk in comparison to that between ESG measures and upside idiosyncratic risk?
    a. Is there a relationship between ESG measures and idiosyncratic risk?
    b. Is there a relationship between ESG measures and downside idiosyncratic risk?
    c. Is there a relationship between ESG measures and upside idiosyncratic risk?
    d. How do these relationships, or lack thereof, compare with each other?
2. Does idiosyncratic risk mediate the relationship between ESG measures and market returns?
    a. Is there a relationship between ESG measures and market returns?
    b. Is there a relationship between ESG measures and the three idiosyncratic risks?
    c. Does the relationship between ESG measures and idiosyncratic risk mediate the relationship between ESG measures and market returns?

The hypotheses developed to address these research questions are described in the section eight *Empirical methodology*. The answers to these research questions are contained in the chapters 9 and 10.

1.3 Delimitations

The analyses in this thesis does not attempt to answer the question of causality, which is not necessarily follow from statistically significant correlation coefficients or regression coefficients. Analysis of correlation can provide evidence of association. Analysis of correlation can provide evidence of a relationship but not a causal relationship (Urdan, 2010).

The inferences made based on the results of the analyses of particular dimensions of ESG measures, market returns, and idiosyncratic risk cannot automatically or easily be applied to the broader concepts of CSR, CFP, and risk or other dimensions contained within these concepts.

Furthermore, results may vary depending on the methodology used to calculate variables used to define ESG measures, market returns, and idiosyncratic risk. Significant considerations may be that the empirical analyses here are limited to one asset pricing model: the Fama French five factor model (Fama & French, 2015, 2016). The analyses here also only use the Asset4 ESG database. There was evidence found that suggested that the Asset4, KLD, and Bloomberg Sustainability ESG measures lack of convergence in ESG measurement and coincide in neither distribution nor risk (Dorfleitner, Halbritter, & Nguyen, 2015).

The analyses in this thesis also only examine those European based firms included in the Asset4 ESG database.

1.4 Structure

Following the introduction, the second part of this thesis begins by exploring the many dimensions that fall under the terms of CSP, CFP, and risk. This is motivated by the methodological issues that may have arisen in empirical research due to the conflation of dimensions despite their distinct nature. The concepts referred to by their respective umbrella terms share different relationships between each other. What exactly is meant by the claim "CSP reduces risk and increases CFP?" Which CSP, which risk, and which CFP? It is essential to define our variables of interest.

The third part contains an analysis of the theoretical arguments offered regarding the relationships between CSP, CFP, and risk. The third part contains a review of the results of relevant empirical studies regarding these relationships.

The fourth part contains a description of the data used in the empirical analyses and the process of their collection. A description of the empirical methodology follows. After this follows a presentation of the results and discussion.

The final section *Conclusion* contains a summary of the results in relation to our research questions and concludes the thesis.

2 Corporate social responsibility and performance

This section presents the concept of CSR and CSP and the complications that arise from their multidimensional and contested nature. The section concludes with determination of the particular measure chosen for empirical analysis.

2.1 Definitions and discussion

This thesis assumes the definition of corporate social responsibility ("CSR") offered by Gord & Moon (2011). They describe CSR as a management idea and academic 'cluster concept' that broadly refers to:

1. The expectation that business is both responsible to society (i.e., accountability) (Bowen, 1953; Carroll, 1979) and for society (i.e., compensating for negative externalities and contributing to social welfare) (Crouch 2006, Arrow 1974);
2. The expectation that business conducts itself in a responsible fashion (Carroll, 1979); and
3. The expectation that business manages the corporation-society interface through the enhancement of stakeholder relationships (Barnett 2007; Gond and Matten 2007; Freeman 1984).

One of its essential characteristic is that it is "contextual, dynamic, and overlapping". It is contextual because the meaning and practices of CSR differ from one context to another. These contexts include country, culture, and scope of analysis (i.e., institutional, organisational, and individual). It is dynamic because it changes with time, within each particular context, as societal mores and demands on business change. It is overlapping because the boundaries of the field are porous, frequently encroaching upon other fields and vice versa, such as business ethics, law, labour relations, political economics, diversity studies, and critical management studies. (Gord & Moon, 2011)

Furthermore, it is a member of the family of "essentially contested concepts". These are a group of concepts which are chiefly characterised by the inclusion of contestation of their meaning as an essential part of their existence, largely due to their aspirational nature and the lack of any authority to settle the contestation (i.e., to choose one true meaning from among the many competing meanings). Another example often given is the concept of freedom. There will never be an authoritative meaning assigned to the concept of freedom because contestation is an essential part of its being (Gord & Moon, 2011).

It is therefore necessary to recognize that there is no authoritative definition of CSR to which one can appeal. Furthermore, even in the case that a definition is be settled upon, there is no authoritative measure or proxy for that particular definition of CSR. Comparisons of CSR across industry, culture, country, and time are therefore complicated because the meanings attributed to CSR and judgments of firm activity related to CSR change across these vectors. This also complicates any study of CSR and particularly empirical research. Some authors argue that the “business case” for CSR is either counterproductive to achieving its goals (e.g., in comparison to normative arguments) (Noon 2007) or not as relevant as assumed because CSR is as much a responsibility thrust upon business as it is an elective opportunity (e.g., political CSR) (Scherer & Palazzo 2008)

Following from the definition of CSR assumed, corporate social performance (“CSP”) is defined as a firm’s success in meeting the expectations that business is responsible to and for society, that it conducts itself in a responsible fashion, and that it manages the corporation-society interface.

This definition also room for contestation and requires clarification. This thesis defines CSP as a firm’s quantified relative performance on an range of CSR related activities classified into distinct environmental, social, and governance categories (ESG measures).

A number of specialized rating institutions research, determine, and offer ESG measures, some of the most important providers being ASSET4 by Thomson Reuters, Ethical Investment Research Service (EURUS), Kinder Lydenberg Domini & Co. (KLD) by MSCI, Sustainability Asset Management Group (SAM), Bloomberg Sustainability (Dorfleitner et al., 2015) and Sustainalytics. An analysis of the ASSET4, KLD, and Bloomberg Sustainability ESG measures suggested a lack of convergence in ESG measurement. The three ESG measures coincided in neither distribution nor risk (Dorfleitner et al., 2015).

2.2 Selection of measure

ESG measures were chosen for study in this thesis to facilitate statistical analysis and testing of relationships between variables of interest. They allow for analysis of the strength of associations, not possible when using binary distinctions (e.g., responsible or not responsible). Relative performance can then to compared to relative risk and return performance, the relationships between which are the subject of the theories tested in this thesis.

3 Corporate financial performance

This section presents the concept of CFP and a comparison of accounting-based measures and market-based measures. The section concludes with the selection of the particular measure chosen for empirical analysis.

3.1 Definitions and discussion

Orlitzky, Schmidt, & Rynes (2003) distinguish between three subdivisions in the operationalization of corporate financial performance ("CFP") in their meta-analysis of the research regarding the effects of CSP on CFP: market-based measures, accounting-based measures, and perceptual measures.

Market-based measures such as returns which depend on the stock market participants to determine a firm's stock price and market value. Accounting-based measures such the firm's return on assets (ROA), return on equity (ROE), or earnings per share (EPS) which reflect internal profitability and reflect internal decision making and managerial performance rather than external market response to organizational actions. Perceptual measures gather subjective estimates from respondents of firm's qualities such as 'soundness of financial position' and 'wise use to corporate assets' (Orlitzky et al. 2003).

Accounting-based measures of CSP are primarily designed to capture a firm's profitability with many variations in their construction to facilitate comparison of firm performance across time and cross sectionally. Profits can be presented as a proportion of sales (net margin), assets (ROA), or invested capital (ROIC), or divided equally amongst shareholders (EPS), or adjusted for non operating and one time events that are unlikely to reoccur (NOPAT), or adjusted for risk (EVA). These measures are primarily historical accounts of profitability and are thus sometimes referred to as "backward looking". However, those accounting profitability measures that adjust for risk using the weighted average cost of capital (WACC) are consequently incorporating forward looking market information via the cost of equity (i.e., market beta) and cost of debt (Petersen & Plenborg, 2012).

In contrast, market-based measures are primarily designed to capture the increase in shareholder value as a result of holding a firm's equity via any increase in firm value and dividends, or other forms of payments, received. These measures are primarily reflections of changes in the market expectation of a firm's future cash flows and risk (i.e., the basis of firm value) and are thus sometimes referred to as "forward-looking" (Petersen & Plenborg, 2012). This thesis is primarily concerned with market-based measures.

The most basic form of market-based measure is the rate of return to an investor of holding an asset over some holding period Δt, defined as:

$$r_{t,t+\Delta t} = \frac{D_{t,t+\Delta t} + P_{t,t+\Delta t} - P_t}{P_t} = \frac{D_{t,t+\Delta t} + P_{t,t+\Delta t}}{P_t} - 1$$

Where P_t and $P_{t,t+\Delta t}$ denote the price of the asset at the beginning and the end of the holding period and $D_{t+\Delta t}$ denotes the cash dividend provided to the holder of the asset over the same period.

More sophisticated market-based measures subsequently adjust for risk, which is the subject of the following section of this thesis.

3.2 Selection of measure

The efficient market hypothesis implies that at any point in time a company's stock price should reflect all available information. Market returns should therefore take into account the historical profitability information present in accounting-based measures but also all other information and considerations. Market returns were chosen therefore primarily because they contain greater informational content: the market's full expectations and predictions of future performance based on all available information.

In addition, in the case that some of the effects of CSP on CFP are long term in nature, they would not appear in contemporaneous accounting measures or those in the years immediately following. In contrast, a stock's price should reflect the expected long-term effects on profitability because they incorporate forecasted cash flows into perpetuity.

The performance of firms was chosen instead of the performance of mutual funds in order to avoid the issue related to the additional variable mutual fund manager. This avoids the possibility of attributing superior market returns to firm CSP when in fact they are attributable to the skills of a portfolio manager (Kempf & Osthoff, 2007).

4 Risk

This section presents the concept of risk, its different interpretations, and the interrelated measures related available. The section concludes with the selection of the particular risk measures chosen for empirical analysis.

4.1 Summary

The concept of risk is essential to many fields. The definition varies both between fields and within each field as well. It can refer to several different categories of concepts, each containing different nuances and proxies used to measure them.

In very broad and colloquial terms, risk measures can be organized into the following three categories:

1. Risk as the chance of a bad event occurring.
2. Risk as uncertainty in outcome.
3. Risk as exposure to the variance of some other thing.

Firstly, risk may refer to the probability of the occurrence of an adverse event. This is what is meant by the term default risk. "An issuer not delivering the promised payments is said to default on the bond. The risk that this may happen to referred to as default risk or credit risk" (Munk, 2016, p. 139). In this case, risk is expressed as an estimated probability of that adverse event occurring (e.g., the default risk of this particular customer is 5%). A similar concept is used when we speak of risk of death, fire risk, or earthquake risk. It is common in discussions regarding insurance and the probability of the occurrence of the triggering event described in a policy.

Secondly, speaking from within in the field of investments, Munk (2016) introduces the concept of risk as follows: "When making your investment, you might have a good idea about the future price and dividends of the asset, but in general you cannot know them for sure. Therefore, the return you obtain is uncertain or, in other words, risky until the end of the investment period." In this definition, risk refers to the uncertainty of the eventual value of some observable future phenomenon. Risk is uncertainty of outcome. Following the example used in the quotation, these phenomena are the future price of an asset and the dividends it provides, which together when compared to the purchase price of an asset generates an investor's return. In this case, risk is expressed as an estimated range of the possible values that the future phenomena may take, along with accompanying probabilities (i.e., a probability distribution of a random variable). The most common example is variance, or its standardized form standard deviation.

Switching to the language of statistics, risk of this second category can be described as the amount of dispersion around a measure of central tendency (Urdan, 2010).

Finally, using the language of statistical regression, risk can refer to the amount of variance in an independent variable of interest that is driven by the variance in a second explanatory variable. This is referred to as the unstandardized regression coefficient. This measure is closely related to the correlation coefficient between the independent and dependent variable. It is determined by transforming the correlation coefficient into the scales of measurement of the two variables (Urdan, 2010).

Often, this risk is referred to as the "*x* risk" or "exposure to *x*" in which *x* would be substituted with the explanatory variable. An example is market risk, measured using the market-beta, in the CAPM. Market-beta refers to the amount of variance in the returns of an asset driven by the returns of the market portfolio and is therefore the regression coefficient of asset returns regressed on market returns. It is not the total variance of market returns around some measure of central tendency (e.g., mean), as measured by variance, but only that variable uniquely attributable to the variance of market returns.

The following sections describe in more detail the measures that are discussed in this thesis.

4.2 Variance and volatility: risk as uncertainty

Related to the second category, the most basic risk measures utilized to quantify the uncertainty regarding investment returns are the variance and the standard deviation of the returns (Munk, 2016). Furthermore, the standard deviation of returns is commonly referred to as volatility.

Firstly, the expected rate of return is defined as the probability-weighted average of the possible rates of return, as follows (Munk, 2016):

$$\mathrm{E}[r] = \sum_{s=1}^{S} p_s r_s .$$

Often the expected return is estimated using the average realized return over some preceding historical period. Past patterns are assumed to contain information useful for predicting the future (Munk, 2016). E.g., if the average daily log return for an asset over the last year have been 0.05%, then the expected return tomorrow is 0.05%.

Around this expected value we can then calculate the variance, defined as the sum of squared deviations from the expected value, weighted by the probability (Munk, 2016), as follows:

$$\mathrm{Var}[r] = \mathrm{E}\left[(r - \mathrm{E}[r])^2\right] = \sum_{s=1}^{S} p_s (r_s - \mathrm{E}[r])^2.$$

The standard deviation of returns is simple the square root of the variance calculated above (Munk, 2016), as follows:

$$\mathrm{Std}[r] = \sqrt{\mathrm{Var}[r]}.$$

The standard deviation includes both positive and negative deviations from the expected value but it can be argued that only the negative deviations (e.g., downside risk) should be included in a risk measure (Munk, 2016) because there are what concern an investor. In case, the lower partial standard deviation can be computed as just the standard deviation using only the realizations below some baseline (e.g., the expected value, the risk-free rate, (Munk, 2016) or zero). Equivalently, the upper partial standard deviation would include only those realizations above the baseline. Analogous downside and upside measures related to the variance can be computed and are referred to as semivariance Washer & Johnson (2013).

Related to the measures that consider dispersion around a central tendency are skewness and kurtosis, which describe the shape of a probability distribution (Munk, 2016). Skewness is a measure of the asymmetry of a variable's distribution. The skewness is zero in the case of any normal (i.e., symmetrical) distribution. If the distribution leans to the left so that more of than half of the probability mass is above the mode, the skew is positive. Conversely, if the distribution leans to the right, the skew is negative.

Kurtosis is the standardized four moment. A distribution with a positive kurtosis will have "fatter tails" than the normal distribution. In such a case there is higher than normal probability of large positive and large negative return realizations.

Investors may be particularly concerned about the possibility of highly negative returns. Value at Risk and expected shortfall are two common measures that attempt to quantify these risks, both focusing on the left tail of the probability distribution (Munk, 2016).

4.3 Market-beta and factor betas: risk as exposure

The concept of risk in the field of investments changes significantly when the perspective of analysis is broadened from that of the single asset to an investor's portfolio of assets.

When assessing the risk of any particular asset, it can be argued that investors are and should be more interested in the contribution of each individual asset to the overall risk of the portfolio (Munk, 2016).

The Capital Asset Pricing Model (CAPM) is the most famous of the models of equilibrium prices of financial assets and was derived by Treynor (1961), Sharpe (1964), Lintner (1965), and Mossin (1966) (Munk, 2016). The following relationship regarding the expected excess returns of any risk asset, when the market is in equilibrium under particular assumptions, (Munk, 2016):

$$\mathrm{E}[r_p] - r_f = \beta_p(\mathrm{E}[r_m] - r_f),$$

Where β_i is the market-beta of asset *i*, defined as:

$$\beta_i = \frac{\mathrm{Cov}[r_i, r_m]}{\mathrm{Var}[r_m]}$$

The CAPM dictates that the expected return above and beyond the risk-free rate (i.e., the risk premium) on any risky asset is the product of the market-beta of the asset and market risk premium. The market-beta is therefore the correct risk measure for each individual asset. The standard deviation or variance of the returns of an asset are irrelevant, and neither is the covariance of the asset return with any other variable than the market return (Munk, 2016). If all other risk can be diversified away, then no investor will be rewarded for being exposed to it, and need not concern themselves with it.

If the CAPM does not hold however, the difference between the expected excess return and the realized excess return is known as the asset's Jensen's alpha, simply "alpha", or abnormal returns. An abnormal return is any return greater than that return expected based on the priced risks taken on (Munk 2016). It is defined as follows:

$$\alpha_i = \mathrm{E}[r_i] - r_f - \beta_i\left(\mathrm{E}[r_m] - r_f\right).$$

The CAPM and other single index models assume that a single factor is the source of the common variation across risky assets. This was considered restrictive and empirical studies suggest that more than one risk factor is priced (Munk 2016). This led to multi-factor models that incorporate additional risk factors that explain some of the variation remaining after consideration of the market-beta.

The well known Fama-French three-factor model (Fama & French, 1992) included the following factors: *market* defined as the return on a broad stock market index, *small-minus-big* (e.g., *SMB*) defined as the return on a portfolio of small companies minus the return on a portfolio of large companies, and *high-minus-low* (i.e., *HML*) defined as the return on a portfolio of stocks issues by firms with a high book-to-market value (i.e., value stocks) minus the return on a portfolio of stocks issued by firms with low book-to-market value (i.e., growth stocks).

This model with anomalies observed in the market which couldn't be explained by the CAPM. After taking into consideration market-beta, stocks in small companies were found to provide higher returns than stocks in large companies and value stocks tend to provide high return than growth stocks. The model assumes the excess returns satisfy the following:

$$r_i - r_f = \alpha_i + \beta_{i,m}\left(r_m - r_f\right) + \beta_{i,SMB}\text{SMB} + \beta_{i,HML}\text{HML} + e_i,$$

Multifactor models are flexible. Carhart (1997) extended the model by adding a fourth momentum factor defined as the difference between the future return on a portfolio of recent \winners" and the future return on a portfolio of recent losers. Fama & French (2015, 2016) added *profitability* and *investment* to the old model to create the Fama French five-factor model.

4.4 Idiosyncratic risk: risk as variance after controlling for exposures

Idiosyncratic risk is also referred to non-systematic, idiosyncratic, asset-specific, firm-specific, or diversifiable risk (Munk, 2016). In the context of a portfolio of assets, all of the covariance is due to the common variation with the market and the idiosyncratic risks are uncorrelated with the market and uncorrelated across assets (Munk, 2016).

Idiosyncratic risk is estimated by forming a time series of residuals from the application of an asset pricing model and taking the same variance of the time series (Munk, 2016). In the case of CAPM model:

$$r_p - r_f = \alpha_p + \beta_p\left(r_m - r_f\right) + e_p,$$

$$\hat{e}_{it} = r_{it} - r_{ft} - \left[\hat{\alpha}_i + \hat{\beta}_i\left(r_{mt} - r_{ft}\right)\right],$$

The estimate of $\text{Var}[e_i]$ is the sample variance of the time series above. The variance of returns can be decomposed into a systematic or market risk component (beta) and idiosyncratic risk component (error term) as follows:

$$\text{Var}[r_i] = \beta_i^2\,\text{Var}[r_m] + \text{Var}[e_i],$$

Idiosyncratic volatility is the square root of $\text{Var}[e_i]$.

Bodie (2014) similarly presents this decomposition and describes e_i, idiosyncratic risk, as the term measuring "firm-specific surprise" that "is independent of shocks to the common factor that affect the entire economy".

In a large balanced portfolio of assets, the non-systematic risk is small, in fact often smaller than the non-systematic risks of the assets comprising the portfolio, and it is reduced by diversification as you increase the number of assets in the portfolio because the non-systemic return component will be positive for some assets and negative for others. Theoretically, if an investor had an infinite number of assets, they could diversity away all non-systematic risk (Munk, 2016).

Therefore, by forming sufficiently large portfolios with small weights in each asset, investors can diversity away the asset-specific risk, but cannot diversity away the market-wide risk, which is captured by the covariances across assets (Munk, 2016).

Berstein (1999, p. 136) expressed the same point elegantly when he wrote that the "the power of diversification obliterates the individual attributes of the stocks, and more than 90 percent of the portfolio's variability is explained by the index."

The process of decomposition of total variance described above can easily extended to include additional factors. Each factor in addition to the market risk premium $(r_t - r_f)$ will receive its own estimated estimate factor beta $\hat{\beta}$ and the idiosyncratic variance would again be calculated as the simple variance of the time series of residuals from the model.

4.5: Selection of measure

This thesis is not concerned with determining how much risk each firm's stock contributes to a fully diversified portfolio (i.e., market-beta) but rather aims to use that risk measure that best captures the effect of shocks to firm value that can be potentially be influenced by CSP.

Idiosyncratic risk is that variance in an asset's return that is not related to shocks to the market and other factors common to all assets (i.e., firm-specific shocks). The types of adverse events, and consequent negative shocks to firm value, that CSP is predicted to reduce the likelihood of (i.e., risk mitigation theories such as Sharfman & Fernando (2008)) or the severity of (i.e., insurance effect of Godfrey (2005)) are firm specific. I suggest that it is for this reason that idiosyncratic risk should be most affected by CSP. CSP should reduce the incidence of and magnitude of firm-specific adverse events, thus reducing idiosyncratic risk.

Luo & Bhattacharya (2009, p. 209) write that "it is reasonable to believe that the relationship between CSP and firm-idiosyncratic risk is stronger than the relationship between CSP and systematic risk." Sassen et al. (2016) similarly speculate that the lack statistically significant evidence of a relationship between aggregate CSP and systematic risk, in contrast to idiosyncratic and total risk, may be due to the fact that systematic risk is driven industry-specific factor as opposed to firm-specific factors. Individual CSP is therefore assumed to be a firm-specific characteristic to which idiosyncratic risk is more responsive.

Bouslah et al. (2013) relegate systematic risk to a footnote. They similarly state that CSP is likely to affect idiosyncratic risk because the implications of CSP practices such as employee commitment, lawsuits, strikes, fines, reputational risk and boycotts are primarily firm-specific in nature.

Furthermore, CSP may have a different affect on downside and upside idiosyncratic risk. The insurance effect in particular theorizes CSP as providing protection of firm value, as opposed to creation of firm value. The risk mitigation theories presents CSP in a similar role reducing downside shocks. These effects may not have the same effect on a firm's upside potential however. Their role is limited limiting downside risk, not upside potential.

CSP could therefore have a different effect on the downside and upside components of *idio* risk. Fombrun, Gardberg, & Barnett (2000) theorizes that CSP can help firms cope with bi-directional risk by firstly generating reputational gains that improve a company's ability to attract resources, enhance performance and competitive advantage and secondly also mitigating the risk of reputational losses due to alienating stakeholders.

A focus of this thesis is therefore also on the downside and upside components of *idio* risk, which may be differently effected by CSP. This is done by examining direct measures of downside and upside idiosyncratic volatility and also examining the related measure of idiosyncratic skewness.

5 Theories regarding relationships

The purpose of this section is to explain and discuss the theories regarding the relationships between our concepts of interest.

5.1 The relationship between CSP and CFP

Orlitzky et al. (2003) summarize that the following theories concerning the overall CSP-CFP relationship suggests a positive relationship, all of which predict a positive relationship.

Instrumental stakeholder theory (Clarkson 1995; Cornell & Shapiro 1987; Donaldson & Preston 1995; Freeman 1984; Mitchell, Agle, & Wood 1997) argues that the satisfaction of various stakeholder groups is instrumental for a firm's financial performance (Donaldson & Preston 1995; Jones 1995). Stakeholder-agency theory argues that the negotiation and contracting processes of stakeholder-management relationships serve as monitoring and enforcement mechanisms to prevent align manager and firm financial goals. In addition, this balancing increases the efficiency of a firm's adaptation to external demands.

Secondly, firm-as-contract analysis (Freeman and Even 1990) argues that high firm performance results not only from the separate satisfaction of bilateral relationships but also from the simultaneous coordination and prioritization of multilateral stakeholder interests. Deriving from instrumental stakeholder theory but also referred to as the 'good management theory (Waddock and Graves 1997), high CSP increases a firm's competitive advantage by weighing and addressing the claims of various stakeholder groups in a fair, rational manner.

Thirdly, the causality of the relationship is also theorized to run in the opposite direction. The slack resources theory argues that prior high levels of CSP may provide the slack resources necessary to engage in CSR and responsiveness (Ullmann 1985; Waddock & Graves 1997). Orlitzky et al. (2003) state that they believe both instrumental stakeholder theory and slack resources theory to be accurate and therefore CSP and CFP are related reciprocally (i.e., a virtuous cycle).

Orlitzky et al. (2003) summarize the theory arguing that this relationship is mediated by a number of underlying mechanisms. They divide these into internal benefits and external benefits. CSP increase internal managerial competencies, contributes to firm knowledge about the market, social, political, technology, and other environments, and thus enhances organizational efficiency. Externally, CSP helps build a positive reputation and general goodwill with its external stakeholders.

Orlitzky et al. (2003) provides the following summation of the internal benefits:

1. New competencies and resources manifested in a firm's culture, technology, structure, and human resources (Barney 1991; Russo and Fouts 1997; Wernerfelt 1984);
2. Managerial competencies from significant employee involvement, organization-wide coordination, and managerial style (Shrivastava 1995); and
3. Scanning skills, processes, and information systems, which increase the organization's preparedness for external changes, turbulence, and crises (Russo and Fouts 1997).

Orlitzky et al. (2003) provides the following summation of the external benefits deriving from the external stakeholder reputation and goodwill stance:

1. Positive image with customers, investors, bankers, and suppliers (Fombrun and Shanley 1990);
2. Improved relations with bankers and investors and thus facilitate their access to capital (Spicer 1978);
3. The ability to attract better employees (Greening & Turban 1997, 2000); and
4. Current employees' satisfaction and commitment (Davis 1973; McGuire, Sundgren, & Schneeweis 1988; Waddock & Graves 1997).

Fombrun et al. (2000) also theorize the following benefits:

5. Enhancement of the trust between existing partners by increasing familiarity and social integration;
6. New partnerships outside of direct business linkages which may spur sales opportunities;
7. A favourable relationship and enhanced perceived legitimacy, particularly aboard, with legislators and regulators;
8. The endorsement of activist groups which may influence consumers and increase sales;
9. Increase legitimacy with local communities; and
10. Favourable coverage from the media.

Orlitzky et al. (2003) note a difference in the strength of the relationships between CSP and the different operationalizations of CSP but does not offer an in depth theoretical explanation. This implicitly assumes that the effects of CSP should be similar on market vs accounting measures.

In contrast, Hong & Kacperzyk (2009) theorize that social norms constrain investors leading them to avoid sin stocks (i.e., publicly traded firms operating in controversial industries such as production of alcohol, tobacco, and gaming). They argue that sin stocks should therefore be cheaper than other stocks (i.e., outperform comparables) because, from the work of Merton (1987) on neglected stocks, the neglect of sin stocks leads to the prices of those stocks being depressed relative to their fundamental values due to limited risk sharing. Secondly, in this case, the CAPM no longer holds and idiosyncratic risk, in addition to market-beta, is a relevant factor for pricing. Sin stocks may have higher litigation risk due to their products, which increases idiosyncratic risk and consequently expected returns. Finally, investors may be simply be irrationally undervaluing firms in sinful industries.

Galema, Plantinga, & Scholtens (2008) similarly describe theoretical explanations for the non-existence or negative relation between CSP and market returns as due to the common basic cause of excess demand for high CSP stocks and a shortage of demand for low CSP shows that leads to overpricing of the high CSP stocks and underpricing of low CSP stocks. In addition to those arguments based on Merton (1987) used by Hong & Kacperzyk (2009), they reference theoretical models based on difference in investor preferences regarding non-financial performance characteristics (Heinkel et al. 2001; Fama & French 2007; Dam 2008).

In addition, assuming that that a sizable proportion of investors exhibit multi-attribute utility functions (Bollen, 2007) (i.e., they derive utility from both returns and CSP), then it may appear that they are overpaying for high CSP stocks when analyzing the situation from the perspective of only risk and return because investors are in fact maximizing utility by investing for both return *and* CSP. They may be effectively very rationally paying for CSP to maximize utility.

It is therefore possible that CSP leads to both higher CFP and lower CFP via difference mechanisms and these impacts are registered differently depending on how the concept of CFP is operationalized. CSP may increase internal firm profitability and therefore accounting measures such as ROA. However, the desirability of the social impact created by CSP may in and of itself be valuable to investors, leading them to drive up the price of the high-CSP stocks, drive down the price of neglected low-CSP stocks, and consequently drive down the relative market returns of high-CSP stocks.

5.3 The relationship between CSP and risk

The common theoretical prediction in this field is that there is a negative relationship between CSP and firm risk (Orlitzky & Benjamin 2001).

Furthermore, many of the theories concerning the particular underlying mechanisms can be grouped into the following three broad categories.

1. Risk mitigation: a decreased likelihood of negative events occurring;
2. Insurance effect: a decreased severity of punishment in event of negative events; and
3. Reduced news hoarding: a decreased amount of negative news hoarding.

Firstly, CSP is often predicted to operate as a risk mitigation strategy. It decreases the likelihood of negative shocks to forecasted cash flows and firm value from exposures of controversial irresponsible behaviour because the firm engages in less irresponsible behaviour (e.g., Sharfman & Fernando 2008; Jo & No 2012).

Secondly, in the event that a negative shock does occur, past CSP blunts its impact on firm value in a manner akin to an insurance policy. This is denoted the "insurance effect". (Godfrey 2005; Godfrey, Merrill, & Hansen 2009; Fombrun et al. 2000)

Thirdly, CSP creates an environment of integrity and transparency in corporate governance. This environment discourages bad news hoarding behaviour by management. This in turns decreases the likelihood of a release of a large amount of accumulated bad news and consequently stock price crash risk (Kim, Li, & Li 2014).

In addition, Orlitzky & Benjamin (2001) present several theories that do not fit in the categories above and which I observed in my literature review to be less commonly evoked. Firstly, low investment in CSP may be interpreted as a lack of management skills and thus cause potential investors perceive low CSP firms as riskier (Alexander & Buchholz, 1978; McGuire et al., 1988; Spicer, 1978). Low CSP firms may also be restricted in their access to market capital due to their exclusion from the potential investment universe via the mechanism of investment screens (McGuire et al., 1988). Finally, CSP may lead to tenuous long-term learning processes in interorganizational cooperation and thus enables the firm to lower transaction costs (Coase, 1937; Hill, 1990;Williamson, 1975, 1985).

In contrast, a positive relationship is less commonly predicted because CSP may be duplicisious, self-serving behaviour on the part of management. Managerial opportunism theory (Preston & O'Bannon, 1997) argues that CSR activities are private benefits to managers and a form of entrenchment strategy. CSP may also be used to cover up risky corporate misbehaviour, as is hypothesized to be the motivation behind the significant philanthropic activities of Enron (Hemingway & Maclagan, 2004).

5.3.1 Risk mitigation

Orlitzky & Benjamin (2001) broadly argue that CSP reduces firm risk from the perspective of instrumental stakeholder theory (T. Donaldson & Preston, 1995; Jones, 1995) or good management theory because a firm's disregard of implicit stakeholder claims may lead to uncertain future explicit claims. Greater attentiveness of stakeholder concerns reduces the probability of legal proceedings and regulatory intervention. In addition, high CSP firms may have incorporated organization principles that are surprise avoiding (King, 1995; cf. Frederick, 1995).

McGuire et al (1988) argue that investors may forecast an increase in firm costs such as government fines and lawsuits due to low CSP, some so severe that they may threaten a firm's very existence. The authors use the example of those filed against pharmaceutical, chemical, and asbestos firms.

Boutin-Dufresne and Savaria (2004) argue that firms that adopt a CSR codes of conduct reduce risk of firm specific adverse events that have the potential to significantly affect profitability, such lawsuits and strikes. They argue that the total risk of a firm includes an ethical risk component. Higher CSP would imply lower ethical risk and consequently lower total risk. The authors suggest that ethical risk could be divided into risk subcategories such as environmental risks, product and commercial practices risks, and quality of life in the workplace.

In addition to environmental ('green') performance increasing economic performance via improved resource efficiency, Sharfman & Fernando (2008) propose an additional theoretical perspective of environmental risk management which acts to decrease the perceived riskiness of a firm's cash flows. They argue that higher levels of environmental performance should be viewed by environmental risk management because this mitigates the risk of litigation from regulators and other stakeholders. This reduces the number of potential claimants on future cash flows through fines, settlements, litigation, or compliance costs. Firms that incorporate environmental risk management into their total risk management are rewarded by financial markets via lower costs of financing.

Lee and Faff (2009), argue that high CSP firms are likely to have lower unsystematic (idiosyncratic) risk for reasons such the ability to mitigate sustainability costs and risk, happier, more stable employees, lower fines, stable production levels, and all the other "business-related virtues" bestowed on leading CSP firms.

They later go on to write firms high CSP firms are able to reduce their company specific business risk by adopting a leading CSP strategy. These risks could include adverse events arising from lawsuits, strikes, brand and reputation erosion, and boycotts. These events could materially influence a firm's profitability and overall risk profile.

Jo & Na (2012) reference and build upon the perspective of Sharfman & Fernando (2008) by interpreting the broader concept CSR as a form of risk management by reducing the probabilities of expected financial, social, or environmental crisis that could negatively affect a firm's cash flows.

5.3.2 Insurance effect

The insurance effect of CSP on firm value was first theoretically modelled comprehensively in the work of Godfrey (2005). Expressed succinctly, "CSR-based moral capital creates value *if* it helps stakeholders attribute the negative event to managerial maladroitness rather than malevolence, and temper their reactions accordingly." (Godfrey et al. 2009, p. 428).

Godfrey (2005) wrote that corporate philanthropic activities can creative moral capital when they are positively morally evaluated by stakeholders, determined by the alignment of the activity with the ethical values of those particular stakeholders and their perception that the activity is genuine as opposed to designed to ingratiate the firm. Positive moral capital acts as insurance when it protects relational wealth against loss by mitigating negative stakeholder assessments and related sanctions when bad acts do occur. Relational wealth consists of relationship-based intangible assets such as affective commitment from employees, legitimacy among communities and regulators, trust from suppliers and partners, and the value of the brand among customers. This mitigation occurs through a mechanism related to the doctrine of mens rea. Under this doctrine, two elements must be present for an offence to occur: a bad act and a bad mind. When bad acts occur, the author argues that it is reasonable to assume that stakeholders conduct an evaluation similar to the mens rae doctrine during the process of determining appropriate sanctions. Positive moral capital encourages stakeholders to give the firm the benefit of the doubt regarding intentionality, knowledge, negligence, or recklessness.

Godfrey et al (2009) references earlier work by Fombrun et al. (2000) that theorized a similar model, but without entering deeply into the specifics. They present corporate citizenship as a strategic tool to manage bi-directional risk: creating reputational gains that improve the ability to attract resource, increase performance, and build competitive advantage and also mitigate the risk of "reputational losses that can result from alienating key stakeholders" (Fombrun et al., 2000, p. 85). The increased upside potential is designated the "opportunity platform" and decreased downside risk is designated the "safety net".

Regarding the topic of the threat of legal action from regulators they write that "corporate citizenship activities help to relay such information, aiding in building a corporate atmosphere that not only mitigates the risk of rogue behavior, but also lessens the risk of conviction and imposition of heavy penalties if and when such behavior does occur" (Fombrun et al., 2000).

Independently of Godfrey (2005) but arriving at the same metaphor of insurance, Peloza (2006) argued that the impact of CSP on CFP should be interpreted in an integrated framework taking into account incremental gains such as increased sale and also risk mitigation/insurance of harmful events. The ability for firms to generate value from CSP in the long term is partially due to the potential for it to act as a buffer against negative events via reputation effects. He argues that firms insure virtually all aspects of their operations and CSR can be justified as a purchase of insurance covering a firm's arguable most valuable asset: its reputation.

Peloza (2006) references earlier theoretical work by Bhattacharya and Sen (2001) that argues that CSP can build a reservoir or goodwill that firms can be subsequently drawn upon in time of crisis. They refer to this ability to gain the benefit of the doubt with customers as resilience and argue that consumers are more likely to forgive negative CSR events when firms have long-standing reputations for positive CSP.

More broadly, Peloza (2006) also references the theoretical earlier work of Fombrun et al (2000) as an influence for his integrated model. He related the upside opportunities and downside safety nets in the model of Fombrun et al. (2000) the two forms of financial return to CSP in his own model: incremental gains as rewards and mitigation of consequences from negative firm behaviour.

Finally, insurance and options are related in so far as they can both be utilized for risk management. For example, portfolio managers can use purchase index put options to limit their downside risk (i.e., portfolio insurance") (Hull, 2012)

Husted (2005) developed the notion of CSR as a real option and its implications for risk management. He concludes with the hypothesis that CSP should be negatively related to ex ante downside business risk. He portrays CSR as a method of containment of possible losses, and recommends a shift in focus away from variance to those that focus on downside outcomes.

5.3.3 Reduction in bad news hoarding behaviour

Kim, Li, & Li (2014) present another distinct mechanism through which CSP can lower risk. Their study found evidence that CSP is negatively associated with future stock price crash risk, which is defined using two measures: the conditional skewness of firm-specific weekly returns over the fiscal year and the down-to-up volatility measure (DUVOL) of the crash likelihood (i.e., the natural logarithm of the ratio of the standard deviation in the "down" weeks to the standard deviation in the "up" weeks).

Their model builds on studies that argue that a prominent predictor of stock price crash risk is the managerial tendency to withhold bad news from investors (e.g., Jin and Myers, 2006; Hutton et al., 2009). These studies argue that managers withhold bad news from investors due to career and compensation concerns until the accumulated bad news reaches a tipping point at which point it is all released and results in a stock price crash.

They argue that firms consider increased disclosure as a form of CSP. If higher CSP performance firms extend the same high ethical standard to financial reporting, they are more likely to maintain a higher level of transparency and are less likely to conceal bad news from investors.

6 Empirical literature

The purpose of this section is to present and discuss empirical findings testing the theories described in the sections above.

6.1 The relationship between risk and return

Classical finance theory, as developed by Markowitz (1952), Sharpe (1964), and Lintner (1965) dictates that there is a positive relationship, or trade-off, between risk and expected market returns in equilibrium. Volatility consists of systematic and idiosyncratic risk but investors are predicted to only earn a premium for holding systematic risk because idiosyncratic risk is diversified away in a fully-diversified portfolio. For this reason, idiosyncratic risk is not priced (Frieder & Jiang, 2007).

This implies that if CSP acts to reduce exclusively idiosyncratic risk, as opposed to market-beta, then there would be no increase in market returns. This also implies that if CSP acts primarily but not entirely on idiosyncratic risk, then finding statistically significant evidence of an association may be difficult for this reason.

The single factor CAPM was later extended to multi-factor models, such as the Fama French three-factor model, in which investors are predicted to earn a risk premium on additional risk factors.

In the presence of market frictions where investors have limited access to information, Merton (1987) shows that investors earn a premium for holding idiosyncratic risk because this risk cannot be fully diversified away. This implies that if CSP acts to reduce firm idiosyncratic risk, then this would lead to a decrease in market returns.

However, contrary to the classical finance theory, the results of Ang, Hodrick, Xing, & Zhang (2006) and Ang, Hodrick, Xing, & Zhang (2009) suggest a negative relationship between high idiosyncratic volatility, measured using the CAPM and Fama-French three factor model, and returns. This surprisingly empirical observation has become known as the “idiosyncratic risk puzzle” (Koch 2010) or “idiosyncratic volatility effect” (Ang et al. 2009).

Ang el al (2006) show that differences in opinion measured by analyst dispersion cannot account for the effect. They also rule out exposure to aggregate volatility risk, size, book-to-market, moments and liquidity effects. Ang et al (2009) rule out explanations based on trading or clientele structures, higher moments, and information dissemination. Their results are out-of-sample to earlier US findings and suggest that the effect is not a sample-specific or country-specific effect; the idiosyncratic volatility effect is observed worldwide. They conclude that it is likely that there is some still unspecified underlying economic source for the phenomenon.

This implies that if CSP acts to reduce firm idiosyncratic risk, then this would lead to an increase in market returns via this as of yet unrevealed mechanism.

6.1.1 Downside and upside idiosyncratic risk

Frieder and Jiang (2007) construct measures entitled downside and upside idiosyncratic volatility and analyze their relationship to future stock returns. Downside idiosyncratic volatility is defined as the semi-standard deviation of negative idiosyncratic returns. Upside idiosyncratic volatility is defined as by the semi-standard deviation of positive idiosyncratic returns. Idiosyncratic volatility is defined as the standard deviation of the residuals from the Carhart four-factor model incorporating the market factor, book-to-market (HML) factor, size factor (SMB) factor, and winners minus losers beta (UMD) factor.

They find evidence that suggests that idiosyncratic volatility predicts future returns at both the monthly and quarterly horizons: there is an inverse relationship. Further decomposition finds that downside idiosyncratic volatility fails to predict future returns and that it is upside idiosyncratic volatility drives the inverse relation.

This implies that if CSP acts to reduce upside idiosyncratic risk, then this would lead to an reduce in market returns. There would be no effect on market returns if it acts to reduce downside idiosyncratic risk.

Koch (2010) also decompose idiosyncratic volatility into its downside and upside components. Consistent with previous studies, their results suggest that low idiosyncratic risk stocks yield significantly higher returns than high idiosyncratic risk stocks. In contrast, the results of Koch (2010) suggest find no effect of differentiation between downside and upside idiosyncratic risk. The correlation between idiosyncratic volatility and the downside and upside idiosyncratic volatility measures were found to be high, leading to the author to conclude that the upside and downside measure capture essentially the same risk.

6.2 The relationship between CSR and CFP

Orlitzky et al. (2003) found evidence of a positive association between CSR and CFP in their meta-analysis of 52 studies. The mean observed correlation r_{obs} was .18 with an observed variance of .06. After correction for sampling and measure errors, the true score (corrected) correlation (ρ) was .36 with a variance of .19. The strength of the association was stronger in the case of the social dimension than the environmental dimension. CSP reputation indices were more highly correlated with CFP than other indicators of CSP. CSP was also more highly correlated with accounting-based measures than market-based measures. These results suggest that the operationalization of CSP and CFP moderate the overall positive association.

A more recent meta-analysis of 251 studies published in the period from 1972 through 2007 by Margolis, Elfenbein, & Walsh (2009) found evidence of an overall small positive effect. The mean correlation effect size *r* was .133, the median *r* was .085, and the study size weighted *r* was .105. There was a stronger association between CSP and accounting-based measures of CFP (r = .151) then market-based based measures of CFP (r = .114). A simple vote counting procedure results in 59% of studies revealing an insignificant relationship, 28% a significant positive relationship, and only 2% a negative relationship.

The relationship between prior CSP and subsequent CFP has an average effect size of .152 which indicates that CSP explains approximately 2.23% of the variance in CFP. The relationship between prior CFP and subsequent CSP had an average effect size of .119. The concurrent relationship has an effect size of .112.

A meta-analysis of 42 studies published in the period from 2006 through 2011 by Wang, Dou, & Jia (2016) found evidence of an overall small positive effect. The mean corrected *r* was .0587. There was a stronger association between CSP and accounting-based measures of CFP (corrected r = .0489) than market-based based measures of CFP (correct r = .0378). Perceptual CFP had the highest corrected *r* at .1852.

The relationship between prior CSP and subsequent CFP had an average corrected *r* of .0319. The relationship between prior CFP and subsequent CSP was statistically insignificant with an average corrected *r* of .0096. The concurrent relationship had an average corrected *r* of .0678.

In contrast, Hong & Kacperzyk (2009), which was note included in the meta-analyses described above, found that sin stocks (i.e., those of firms operating in controversial industries) have higher expected returns than otherwise comparable stocks.

Galema et al. (2008), also not included in the meta-analyses, found that CSP impacted stock returns by lowering the book-to-market ratio and not by generating positive abnormal returns.

6.3 The relationship between CSR and risk

The results of the meta-analysis of 18 primary studies by Orlitzky & Benjamin (2001) provide support for the theoretically predicted negative relationship between CSP and risk. The negative association was stronger in the case of accounting risk measures than market risk measures.

The results of studies published following Orlitzky & Benjamin (2001), some of which are discussed below, also generally provide support for the theoretically predicted negative relationship between CSP and risk.

Studies published since then have also added nuance to our understanding of the relationship. Recent insights include the differences between CSP strengths and CSP concerns (Oikonomou et al. 2012; Harjoto & Jo 2014), different dimensions of CSP (Oikonomou et al. 2012; Bouslah et al. 2013; Chang, Kim, & Li 2014; Harjoto & Jo 2015; Sassen et al. 2016), and the role of mediating and moderating variables (Luo & Bhattacharya 2009; Harjoto & Jo 2014; Kim, Li, & Li 2014; Zhang, Xie, & Xu 2016; Lee, M.-T. 2016)

Boutin-Dufresne and Savaria (2004) found a negative relationship between aggregate CSP, measured using the Canadian Social Investment Database, in a sample of Canadian firms during the period from 1995-1999, and idiosyncratic risk measured using the residuals from the CAPM. They did not investigate total or systematic risk.

Bassen, Meyer, & Schlange (2006) found that a negative relationship between aggregate CSP, measured using a custom built scoring methodology, in a sample of diversified utilities firms based in developed markets during 2004 (primarily in the US and Europe), and systematic risk measured using market-beta from the CAPM. They did not investigate total or idiosyncratic risk.

Sharfman and Fernando (2008) found a negative relationship between environmental performance, measured using a custom mix of TRI and KLD data, and systematic risk measured using market-beta from the CAPM. They did not investigate total or idiosyncratic risk.

Luo and Bhattacharya (2009) found a negative relationship between aggregate CSP, measured using *Fortune*'s Most Admired Companies ranking, in a sample of firms during the period from 2002-2003, and idiosyncratic risk measured using the residuals from the Carhart four-factor model. They found the effect to be greater for firms with higher advertising spending. They also found results that support the hypothesis that the simultaneous pursuit of CSP, advertising, and *R*&D increases idiosyncratic risk.

They also found a negative relationship between aggregate CSP and systematic risk measured using market-beta from the Carhart four-factor model. They did not investigate total risk.

Lee and Faff (2009) found a negative relationship between aggregate CSP, measured using leading and lagging portfolios from the Dow Jones Sustainability Index, during the period from 1998-2002, and idiosyncratic risk measured using the residual from the CAPM and a six factor model.

Their results also find that the leading CSP firms do not underperform the market portfolio and their lagging counterparts outperform the market portfolio and the leading portfolio. Further analysis suggested that idiosyncratic risk might be priced by the markets. They conclude that this is evidence that the higher returns for the lagging CSP firms is compensation for higher idiosyncratic risk.

Salama, Hinze, & Hardeck (2011) found a negative relationship between aggregate CSP, measured using community and environmental responsibility (CER) rankings from *Management Today* magazine, in a sample of UK firms during the period from 1994-2006, and systematic risk measured using market-beta from the CAPM. They did not investigate total or idiosyncratic risk.

Oikonomou et al. (2012) argue that corporate socially responsible and corporation socially irresponsible actions are empirically and conceptually distinct constructs. As a result, they decomposed aggregate CSP into CSP strengths and CSP concerns. They found evidence to suggest a weak and negative relationship between aggregated CSP strengths, measured using KLD data, in a sample of US firms, and systematic risk, measured using market-beta from the CAPM. They found evidence to suggest a strong and positive relationship between aggregate CSP concerns and systematic risk, measured using market-beta from the CAPM. They did not investigate total or idiosyncratic risk.

They corroborated this using an analysis of systematic risk using two measures of downside beta, one using the risk free rate as the target return and the second using the mean market rate, resulting in similar results.

The results are more diverse when analysed at the more granular level of individual CSP dimension, as categorized by KLD. In the case of CSP strengths only: there is no statistically significant relationship between systematic risk and any of the dimensions during the entire period. Diversity is positively associated during the high volatility period subsample. Diversity, Employees, and Product dimensions are negatively associated in the low volatility subsample.

In the case of concerns only: Community, Employees, and Environment concerns are positively associated with systematic risk during the entire period. The Employees dimension is positively associated during the high volatility period subsample. The Environment dimension is marginally positively associated as well. The Community dimension is positively associated in the low volatility subsample. The Environment dimension is marginally positively associated as well.

Jo & Na (2012) found that a negative relationship between aggregate CSP, measured using KLD data, in a sample of US firms in controversial industries during the period from 1991-2010, and total risk, measured using the standard deviation of daily stock returns, and systematic risk measured using the market-beta from the CAPM. They did not investigate idiosyncratic risk. They also find the relationship to be more economically and statistically significant in controversial industry firms than in non-controversial industry firms.

Albuquerque, Durnev, & Koskinen (2014) found a negative relationship between aggregate CSP, measured using KLD data, in a sample of US firms during the period from 2003-2011, and systematic risk measured using the market-beta from a three factor model. They did not investigate total or idiosyncratic risk.

Bouslah et al. (2013) presented an examination of the relationship between the individual dimensions of CSP, measured using KLD data, in a sample of US firms during the period from 1991-2007, and total risk, measured using the annualized standard deviation from daily stock returns, and idiosyncratic risk, measured using the residuals from the Carhart four-factor model using the previous year's daily excess returns. The results varied depending on strength vs concern, dimension of CSP, and measure of risk.

Chang et al. (2014) differentiate between CSR activities that target secondary stakeholders, denoting this institutional CSR (ICSR), and those that target primary stakeholders, denoting this technical CSR (TCSR).

They found a negative relationship between ICSR strengths, measured using KLD data, in a sample of US firms during the period from 1995-2009, and systematic risk, measured using the market-beta from the CAPM, and total risk measured using "elative stock volatility to the market. Relative stock volatility to the market is calculated by dividing the 12 month stock volatility by the 12 month CRSP value weighted index volatility. The not find any evidence for a negative relationship between TCSR strengths and the same risk measures.

Harjoto & Jo (2014) found a negative relationship between changes in aggregate CSP, measured using KLD data, in a sample of US firms during the period from 1991-2009, and changes in total risk measured using the standard deviation of monthly stock returns. They did not investigate systematic or idiosyncratic risk.

They also investigate the mediating role of analyst and find statistically significant evidence. They then decompose aggregate CSP into CSP strengths and CSP concerns and find evidence of a mediating role of financial analysts between CSR concerns, but not strengths, and total risk.

Harjoto & Jo (2015) found a negative relationship between changes in CSP, measured using KLD data, in a sample of US firms during the period from 1993-2009, and total risk measured the standard deviation of stock returns. They also found a negative relationship between CSP and both the dispersion of analysts' quarterly earnings forecasts and cost of capital. They did not investigate systematic or idiosyncratic risk.

When decomposed into legal CSP and normative CSP, legal CSP was found to share the same relationships with the risk measures as aggregate CSP, while the opposite was true in the case of normative CSP. They suggest the mechanism to be a form of the risk mitigation effect with emphasis on reputation building.

Sassen et al (2016) found a negative relationship between aggregate CSP, measured using aggregate Asset4 ESG score, in a sample of European firms during the period from 2002 to 2014, and total risk and idiosyncratic risk, but not systematic risk. Total risk is measured using the annualize standard deviation of daily stock returns. Idiosyncratic risk is measures using the annualized standard deviation of the residuals from the Carhart four-factor model applied to daily excess returns over the year. Systematic risk is measured using the market-beta from the CAPM.

CSP was decomposed into social, environment, and corporate governance performance. The found a negative relationship between social performance and all three risk measures. Environmental performance decreased idiosyncratic risk. They found no statistically significant evidence of a relationship between corporate governance performance and any of the risk measures.

6.3.1 The relationship between CSR and downside risk

Research has also conducted recently investigating the relationships between CSP and different downside risk measures.

Kim, Li, & Li (2014) found a negative found a negative relationship between aggregate CSP, measured using KLD data, in a sample of firms during the period from 1995-2009, and future stock price crash risk, measured using both the conditional skewness of firm-specific weekly returns over the fiscal year and the down-to-up volatility (DUVOL) measure of the crash likelihood. They found the relationship to be weaker more pronounced when firms have less effective corporate governance or a lower level of institutional ownership, suggesting the more important role of CSP on risk in the absence of governing mechanisms.

Furthermore, they found evidence of a statistically significant negative association between both risk measures and aggregate CSP strengths and insignificant results when aggregate CSP concerns are analyzed. This suggests CSP strengths drive the overall negative relation between CSP and crash risk.

They did not find any statistically significant evidence when decomposing aggregate CSP into its individual dimensions. They suggest that the relation between CSP and crash risk is therefore mainly driven by the combined effects of CSP across all five dimensions rather than any single dimension.

Zhang, Xie & Xu (2016) similarly found a negative association between CSP and stock price crash risk in a sample of firms in the Chinese market. The negative relationship stronger in the case of state-owned enterprises as opposed to non-state-owned enterprises, and weaker after firms accomplish the split share reform. Lee, M.-T. (2016) similarly found a negative association in a sample of firms in Taiwan.

Diemont, Moore, & Soppe (2016) studies the relationship between the components of CSP, measured using Dutch Sustainability Ratings (DSR) and Sustainayltics databases, and downside equity tail risk defined as the Value at Risk (VaR), at a 1% level, of *idio* returns using the CAPM. The results varied by time period and market.

There was a positive association between the Community component and VaR during the 2008-2011 period and throughout the entire 2003-2011 period in America. The Employees component had a positive association during the 2004-2007 period and a negative association during the 2008-2011 period. The Governance component had a positive association throughout the entire 2003-2011 period in Europe. The Contractors component had a positive association during the 2004-2007 period, a negative association during the 2008-2011 period, and a negative association throughout the entire 2003-2011 period in Europe. The Environment component had a negative association during the 2005-2009 period and a positive association throughout the entire 2003-2011 period in Europe. The Customers component had a negative association during the 2008-2011 period and a negative association throughout the entire 2003-2011 period in America.

Wang, L., Lin, C., Kao, E, & Fung, H (2016) found a negative relationship between CSP, measured using charitable activities, and default risk measured using VaR and expected shortfall measures. They suggest the mechanism to be a higher credit rating, improved employee productivity, and less earnings management. This implies that charitable firms are more productive and enjoy greater financial flexibility and corporate governance.

7 Data

The purpose of this section to provide a description of the data collection process undertaken in order to conduct the empirical analysis.

7.1 Sample selection

Europe was chosen as the market of analysis because of the relatively low amount of research conducted on this market specifically in regards to CSP and risk, in comparison to the US market, in combination with the relatively high emphasis placed on CSR in the market (Sassen et al. 2016). Of the total 14 studies reviewed by Sassen et al. (2016), nine focused exclusively on US market, three on the world market, one on the Canadian market, one on the UK market, and none on the European market. Their literature review focused on those studies published after the meta-analysis of Orlitzky & Benjamin (2001) which found that the US market was the dominant research focus.

Firms in the European market were limited to those covered by Thomas Reuters Asset4 ESG database, designated in the database as "ASSET4 Europe". At the time of retrieval in December 2017, Asset4 Europe consisted of 1,090 firms. The total number of firms globally in the database was 5,933.

The database only contains CSP information for firms in Asset4 Europe since the beginning of the 2002 fiscal year. The number of firms has grown to the current 1,090 from the original 339 firms in 2012. All firm-year observations lacking complete data were eliminated.

7.2 Dependent variables

This section describes the methodology followed in the collection of dependent variables of interest.

Excess log returns

Monthly excess log returns are defined as the excess of the natural logarithm of the monthly returns of the stock of each firm in the sample over the natural logarithm of the monthly risk free rate.

$$r_{it}^{e} = \ln\left(\frac{p_{it}}{p_{i(t-1)}}\right) - \ln\left(r_{ft}\right)$$

Where $p_{it}$is the adjusted price of at the end of month t, $p_{i(t-1)}$ is the adjusted price of at the end of month $(t-1)$ (i.e., at the end of the immediately preceding month), and r_{ft} is the monthly risk-free rate.

Prices were collected from the Thomson Reuters Datastream 5.1 database. They are defined as the Price (Adjusted – Default) data point (Datastream code P). P is defined in Datastream as "the official closing price… The ‘current’ prices taken at the close of market are stored each day. These stored prices are adjusted for subsequent capital actions, and this adjusted figure then becomes the default price…”

Adjusted, as opposed to raw unadjusted prices, were taken in order to arrive at the total return of holding: the sum of the capital gain and any dividends paid during the holding period. to avoid an underestimation of the total returns that accrue to investors which can have a severe impact on cumulative returns over longer investment horizons spanning several years (Brooks, 2014). Which also maintains comparability with the returns of the Fama/French factor portfolios, published by Kenneth French, which include dividends.

Log returns were used to maintain consistency with the majority of academic finance literature, and to take advantage of their time-additive properties. The disadvantage of increased computation of portfolio log returns, in contrast to portfolio simple returns, is not application because portfolios are not the unit of study (Brooks, 2014). Taking logarithms also helps overcome the problem of heteroscedasticity in the regression models used (Brooks, 2014).

The monthly risk-free rate was obtained from Kenneth French’s website and packaged together with the factor portfolio data. Monthly excess log returns are summed to arrive at a set of 10,827 yearly excess log returns, one for year firm-year observation.

7.3 Abnormal returns

Monthly abnormal returns α_i were determined using the Fama French five-factor model (Fama & French, 2015, 2016). The model is defined here as follows:

$$r_{it}^{e} = \alpha_i + \beta_{i,m}\ln\left(r_{mt} - r_{ft}\right) + \beta_{i,SMB}\ln(SMB_t) + \beta_{i,HML}\ln(HML_t) + \beta_{i,RMV}\ln(RMW_t) + \beta_{i,CMA}\ln(CMA_t) + u_{it}$$

Excess returns are modelled as a linear function of five factors: the market, a size factor SMB_t, a style (i.e., value vs growth) factor HML_t, a profitability factor RMW_t, and an investment factor CMA_t.

$r_{mt} - r_{ft}$ is the market risk premium, defined as the return on the European region’s value-weigh portfolio minus the U.S. one month T-bill rate. SMB_t (Small Minus Big) is the average monthly return on the nine small stock portfolios minus the average return on the nine big stock portfolios. HML_t (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios. RMW_t (Robust Minus Weak) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios. CMA_t (Conservative Minus Aggressive) is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios.

For each year and firm in the sample, a multiple linear regression was conducted of the monthly excess log returns (i.e., the independent variable) regressed on the returns of the factor portfolios (i.e., explanatory variables) as described above. The series of five β_i values corresponding to each factor is the exposure of that particular stock in that year to that factor portfolio, and therefore those factors that determine excess returns. They are the series of coefficient estimates on each respective explanatory variable in the regression.

The abnormal returns α_i are the yearly excess returns of the stock that are attributable to that stock's exposure to the factors. Abnormal returns α_i are the intercept estimate of the regression. We conclude with a set of 10,827 α_i values, one for each firm-year observation.

7.4 Idiosyncratic volatility

This thesis adapts the methodology of Sassen et al. (2016), but substitutes monthly returns in place of daily returns and the Fama French five-factor model in place of the Carhart four-factor model, to calculate yearly the idiosyncratic volatility measure $idio_{it}$. Idiosyncratic volatility $idio_{it}$ is defined as the standard deviation of the monthly residuals from the Fama French five-factor model described above. Idiosyncratic volatility is defined as follows:

$$idio_{i,t} = \sqrt{\frac{\sum\left(r_{i,t}^{e} - \widehat{r_{i,t}^{e}}\right)^2}{n-1-k}} = \sqrt{\frac{\sum u_{i,t}{}^2}{n-1-k}}$$

Where k is the degree of freedom adjustment for the number of estimated parameters in the linear regression model (i.e., five). This is equivalent to the standard error of the estimate derived from the linear regression of excess monthly log returns on the monthly returns of the factor portfolios.

Furthermore, this *idio* is annualized using the common adjustment of multiplying by the square root of 12, the number of periods used in its construction in the year, as follows:

$$idio_{i,t}^{a} = \sqrt{\frac{\sum u_{i,t}{}^2}{n-1-k}} \cdot \sqrt{12} = idio_{i,t} \cdot \sqrt{12}$$

Idiosyncratic volatility $idio_{i,t}^{a}$ measures the variance in the stock's return that is not attributable to the variance in the five factors. It measures firm-specific shocks not related to shocks to the market and other factors common to all assets.

7.5 Downside and upside idiosyncratic volatility

In order to investigate whether the relationship between CSP and idiosyncratic volatility differs from the relationship of CSP to idiosyncratic volatility's downside and upside components, this thesis adapts the methodology of Frieder & Jiang (2007) and Koch (2010) to decompose the risk measure. The two components are defined as follows (Frieder & Jiang, 2007):

$$ds.idio_{i,t} = \sqrt{\frac{\sum_{u_{i,t}<0} {u_{i,t}}^2}{n-1-k}}$$

$$up.idio_{i,t} = \sqrt{\frac{\sum_{u_{i,t}\geq 0} {u_{i,t}}^2}{n-1-k}}$$

Furthermore, the same animalization adjustment in this case as to $idio_{it}$:

$$ds.idio_{i,t}^{a} = ds.idio_{i,t} \cdot \sqrt{12}$$

$$up.idio_{i,t}^{a} = up.idio_{i,t} \cdot \sqrt{12}$$

The following relationship between the risk measures holds by construction:

$$\left(idio_{i,t}\right)^2 = \left(ds.idio_{i,t}\right)^2 + \left(up.idio_{i,t}\right)^2$$

$$\left(idio_{i,t}^{a}\right)^2 = \left(ds.idio_{i,t}^{a}\right)^2 + \left(up.idio_{i,t}^{a}\right)^2$$

7.6 Idiosyncratic skewness

In addition to the decomposition above, the third moment of residuals was calculated in order analyze the relationship between CSP and a measure of the asymmetry of their distribution about the mean. Following the methodology of Mitton & Vorkink (2007) and Boyer, Mitton, & Vorkink (2010), idiosyncratic skewness $is_{i,t}$ is defined as:

$$is_{i,t} = \frac{1}{n-2} \cdot \frac{\sum {u_{i,t}}^3}{{idio_{i,t}}^3}$$

The calculation of this explicit risk measure allows for a linear regression of CSP on idiosyncratic skewness $is_{i,t}$ to test the hypothesis that CSP drives upside potential and protects against downside volatility.

The measures described in the sections above are closely related. Idiosyncratic skewness $is_{i,t}$ will be higher in those cases in which upside idiosyncratic volatility $up.idio_{i,t}$ is a greater proportion of total idiosyncratic volatility $idio_{i,t}$ than downside idiosyncratic volatility $ds.idio_{i,t}$. These relations are analogous to those between total variance, downside semivariance, upside semivariance, and skewness, but with the focus shifted from total variance of returns to those firm specific.

Washer & Johnson (2013) wrote that the ratio of downside semivariance to total variance indicates the portion of the variance emanating from the left side of the distribution and that when the distribution has a negative skewness measure, downside semivariance is almost always greater than 50 percent of the of the total variance.

Equivalently, when the ratio of upside semivariance to total variance is greater than 50, one should expect the distribution to have a positive skewness measure.

This is reflected in the sample data collected. We can define the proportion of upside idiosyncratic volatility to total idiosyncratic volatility as simply:

$$up.proportion_{i,t} = \frac{up.idio_{i,t}}{idio_{i,t}}$$

The correlation coefficient of $up.proportion_{i,t}$ and $is_{i,t}$ is a highly statistically significant and strong 0.9819.

7.7 Independent variables: ESG measures

The measures of CSP used in this analysis are the ESG scores provided by the Thomson Reuters Asset4 database, following the methodology of Sassen et al (2016).[1]

The Asset4 ESG data set is built using over 850 binary data points related to sustainability reporting. These data points are aggregated into over 250 ESG ‘Key Performance Indicators” (KPIs), which are combined into 18 category scores that form four pillars of social performance: economic, environmental, social, and corporate governance. A firm can attain any value between 0 and 100 in each pillar. A higher score represents stronger performance in that particular pillar in relation to all rated companies. The four pillars are equally weighted to compute a total ‘Integrated Rating’. The scores in each pillar are normalized using z-scores in order to facilitate benchmarking (Dorfleitner et al., 2015).

The economic pillar was designed to measure “a company’s capacity to generate sustainable growth and a high return on investment through the efficient use of all its resources. It is reflection of a company's overall financial health and its ability to generate long term shareholder value through its use of best

[1] Please note that Thomson Reuters is significantly revising their ESG rating methodology and therefore rebranding and replacing the existing Asset4 ESG scores as “Thomson Reuters ESG Scores” in 2017 (Thomson Reuters, 2017).

management practices." (Thomson Reuters, 2013). The economic pillar is composed of the following three equally weighted categories: client loyalty, performance, and shareholder loyalty (Thomson Reuters, 2017).

The environmental pillar was designed to measure "a company's impact on living and non-living natural systems, including the air, land and water, as well as complete ecosystems. It reflects how well a company uses best management practices to avoid environmental risks and capitalize on environmental opportunities in order to generate long term shareholder value." (Thomson Reuters, 2013). It is composed of the following three equally weighted categories: resource reduction, emission reduction, and product innovation Thomson Reuters, 2017).

The social pillar was designed to measure "a company's capacity to generate trust and loyalty with its workforce, customers and society, through its use of best management practices. It is a reflection of the company's reputation and the health of its license to operate, which are key factors in determining its ability to generate long term shareholder value." (Thomson Reuters, 2013). It is composed of the following seven equally weighted categories: employment quality, health and safety, training and development, diversity and opportunity, human rights, community, and product responsibility Thomson Reuters, 2017.

The corporate governance pillar was designed to measure "a company's systems and processes, which ensure that its board members and executives act in the best interests of its long term shareholders. It reflects a company's capacity, through its use of best management practices, to direct and control its rights and responsibilities through the creation of incentives, as well as checks and balances in order to generate long term shareholder value." (Thomson Reuters, 2013). It is composed of the following five equally weighted categories: board structure, compensation policy, board functions, shareholder rights, and vision and strategy (Thomson Reuters, 2017).

The economic pillar and the total integrate rating are not relevant for purposes of measuring CSP in this analysis. The Datastream codes for the social, corporate governance, and environmental pillars are *SOCSCORE*, *CGVSCORE*, and *ENVSCORE*. Each pillar can take on a value from 0 to 100. Following the methodology of Sassen et al (2016), these individual pillar scores are divided by 100 to arrive at a range from 0 to 1 and equally weighted to arrive at an aggregated measure of CSP designated ESG.

This analysis in this thesis focuses on both the three individual CSP pillars and the aggregated CSP measure ESG because it has been empirically demonstrated that the CSP of individual firms differs across dimensions and the relation of each dimension to each risk measure differs as well (Boulash et al. 2013; Bassen et al. 2006; Diemont et al. 2016; Oikonomou et al. 2012; Harjoto & Jo 2015; Chang & Li 2014). Theoretically, CSP is a multidimensional construct that embodies many dimensions and the expected impacts on different risk measures of each dimension may differ. It would be unreasonable to assume that investors are homogenous with respect to their definitions of each dimensions and its effect on investment decisions (Boulash et al, 2013). Analysis at the aggregate level may therefore obscure the relations of the individual dimensions on risk.

It has also been demonstrated that CSP strengths and CSP concerns differ in their relation to risk measures (Oikonomou et al (2012), Boulash et al, (2013), Harjoto & Jo (2015)). The analysis in the aforementioned studies was aided by the classification of CSP into strength and concern indicators in the KLD rating model of MSCI ESG Research (Dorfleitner et al., 2015). In contrast, Thomson Reuters does not do the same for the Asset4 ESG data set. As a result, this thesis follows the methodology of other studies using the Asset4 ESG data set and no such separation is conducted (Sassen et al 2016).

7.8 Independent variables: control variables

This thesis follows the methodology of previous studies by including several firm characteristics as control variables in the regression analysis in order to better isolate the relationship between the dependent variables of risk measures and the independent variables of interest (i.e., CSP measures) (Sassen et al 2016; Bouslah et al. 2013; Chang et al. 2014; Luo and Bhattacharya 2009; Oikonomou et al. 2012).

Firm size (Size) is defined as the natural logarithm of Total Assets at fiscal year end (Datastream code WC02999) denominated in thousands of USD. Total assets is defined by Datastream as the "sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets." The inclusion of Size controls for the relationship between firm size and risk measures.

Return on assets (ROA) is defined as the Pretax Income for the fiscal year ended (Datastream code WC01401) divided by Total Assets at fiscal year end (Datasteam code WC02999), both denominated in thousands of home currency. Total Assets is defined as in the paragraph above. Pretax Income is defined by Datastream as "all income/loss before any federal, state or local taxes. Extraordinary items reported net of taxes are excluded." The inclusion of ROA controls for the relationship between profitability and risk measures.

Standard deviation of return on assets (SDROA) is defined as the yearly standard deviation of return on assets (ROA), as defined in the paragraph above, over the preceding five years. Its inclusion controls for the relationship between risk measures and profitability volatility as a sign of uncertainty.

Firm leverage (Leverage) is defined as define as Long Term Debt for the fiscal year ended (Datastream code WC03251) divided by Total Assets at fiscal year end (Datastream code WC02999), both denominated in thousands of home currency. Total Assets is defined as in the paragraph regarding Size above. Long term debt is defined by Datastream, as "all interest bearing financial obligations, excluding amounts due within one year. It is shown net of premium or discount." The inclusion of Leverage controls for the relationship between firm capital structure on risk measures.

The market-to-book ratio (MTB) is defined as Market Capitalization at fiscal year end (Datastream code WC08001) divided by Common Equity at fiscal year end (Datastream code WC03501). Market Capitalization is defined by Datastream as "Market Price-Year End * Common Shares Outstanding". Common Equity is defined by Datastream as "common shareholders' investment in a company." MTB was furthermore winsorized at the 0.01 and 0.99 level in order to control for the influence of outliers in the analysis. Several firms had extremely highly negative MTB values due to the effects of negative common equity, specifically due to the inclusion of reserves. The inclusion of MTB controls for the relationship of a different risk measure characteristics of grown and value companies. High MTB is related to value stocks while low MYB is related to growth stocks.

Firm liquidity (Liquidity) is defined as the yearly Turnover by Volume (Datastream code VO) divided by the Common Shares Outstanding at fiscal year end (Datastream code WC05301), both denominated in thousands. Turnover by Volume is defined in Datastream as "the number of shares traded for a stock on a particular day." Common Shares Outstanding is defined by Datastream as "the number of shares outstanding at the company's year end. It is the difference between issued shares and treasury shares." Liquidity was furthermore winsorized at the 0.01 and 0.99 evel in order to control for the influence of outliers in the analysis. The inclusion of Liquidity controls for the relationship between a firm's stock market liquidity and risk measures.

A firm's dividend payment (Div.pay.1) is defined as Dividends Per Share Fiscal (Datastream code WC05110), expressed in actual amounts in the home currency, divided by average price per share during the calendar year. Div.pay is furthermore included in the regression with a time lag of one year. Average share price per share during the calendar year was defined by the average of the month end Unadjusted Price (Datastream code UP), expressed in the home currency. Dividends Per Share Fiscal is defined by Datastream as "the total dividends per share declared during a company's fiscal year. It includes extra dividends declared during the fiscal year but excludes special dividends." Unadjusted Price is defined by Datastream as "the closing price which has not been historically adjusted for bonus and rights issues. This figure therefore represents actual or 'raw' prices as recorded on the day." Div.pay.1 was furthermore winsorized at the 0.01 and 0.99 level in order to control for the influence of outliers in the analysis. The inclusion of Div.pay.1 controls for the relationship between risk measures and the signal expressed by dividend payments of management's perspective of certainty of future earnings. In addition, high expect stock flows are likely to reduce stock volatility.

8 Empirical methodology

8.1 Regression methodology

8.1.1 Determination of model

A dataset that comprises both time series and cross-sectional data is referred to as panel data or longitudinal data. Furthermore, a distinction is drawn between a balanced panel and an unbalanced panel, the difference being that a balanced panel has an equal number of time series observations for each cross-sectional unit, while an unbalanced panel does not (Brooks, 2014).

The simplest method of modelling panel data is a pooled regression in which all independent and dependent variables are "stacked" into responsible single columns containing all of the cross-sectional and time series observations. This method implicitly assumes that the average values of the variables and the relationships between them are constant over time and across all of the cross-sectional units (Brooks, 2014), which may very well not be the case. This would not take into account any common structure in the series of interest (Brooks, 2014), such as a particularly volatile year on the markets.

The two broad classes of panel estimator approach in financial research: fixed effects models and random effects models. These allow the modeler the advantage of being to address a broader range of more complex problems, increasing the number observations, degrees of freedom and thus the power of hypothesis tests, and thirdly the ability to remove the impact of certain forms of omitted variables bias (Brooks, 2014).

In the context of testing the CAPM model, Fama & MacBeth (1973) proposed to handle panel data using a two-step estimation procedure. First, the betas are estimated in separate time series regressions for each firm, and second, for each time period a cross-sectional regression of the excess returns on the betas is conducted. In the second stage, they propose then taking the average of the parameter estimates to conduct hypothesis tests. This approach could be applied in this thesis. However, it is possible to achieve a similar objective using a panel approach (i.e., fixed or random effects) (Brooks, 2014). A panel approach was chosen for this thesis, partially in order to allow more direct comparability with Sassen et al (2016), the only study of CSP and risk using KLD ESG data that could be located during the literature review conducted.

A simple linear regression model can be written as follow (Brooks, 2014):

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

The disturbance term u_{it} can then be decomposed in one of three ways: firstly, into a firm specific effect μ_i and remainder disturbance v_{it}, secondly, into a time specific effect λ_t and v_{it}, or thirdly, utilizing both effects and therefore decomposing into μ_i, λ_t, and v_{it}.

Accordingly, a firm-fixed effects model could be written as follows (Brooks, 2014):

$$y_{it} = \alpha + \beta x_{it} + \mu_i + v_{it}$$

In this model, $\mu_i$is a firm-varying intercept that captures all of the variables that affect the dependent variable cross-sectionally and are constant across time periods (e.g., the industry in which a firm operates, country of headquarters, etc.)

Similarly, a time-fixed effects model could be written as follows (Brooks, 2014):

$$y_{it} = \alpha + \beta x_{it} + \lambda_t + v_{it}$$

In this model, λ_t is a time-varying intercept that captures all of the variables that affect the dependent variable and that vary over time and are constant cross-sectionally across firms (e.g., a change in the regulatory environment or tax rate). These changes could influence the independent variable y, but all firms in the same way (Brooks, 2014).

Finally, the two fixed effects models above could be combined written as the following firm-fixed effects and time-fixed effects model (Brooks, 2014):

$$y_{it} = \alpha + \beta x_{it} + \mu_i + \lambda_t + v_{it}$$

Furthermore, a random effects model approach also includes different intercept terms for each firm, in the case of a firm random effects model, or different intercept terms for each time period, in the case of a time random effects model, or both in the case of a firm random effects and time random effects model.

However, the random effects approach diverges by assuming that, in the case of a firm random effects model, the intercepts for each firm are assumed to arise from a common intercept α, which is constant for all firms and over time, plus a random variable $\epsilon_i$that varies cross-sectionally but is constant over time and measures the random deviation of each firm's intercept term from the "global" intercept term α (Brooks, 2014).

Accordingly, a firm random effects model could be written as follows (Brooks, 2014):

$$y_{it} = \alpha + \beta x_{it} + \omega_{it}, \quad \omega_{it} = \epsilon_i + v_{it}$$

An analogous assumption is made in the case of a time random effects model. It could be written as follows:

$$y_{it} = \alpha + \beta x_{it} + \omega_{it}, \quad \omega_{it} = \epsilon_t + v_{it}$$

Finally, as in the case of the fixed effects approach, the two random effects can be combined in a firm random and time random effects model which can be written as follows:

$$y_{it} = \alpha + \beta x_{it} + \omega_{it}, \ \ \omega_{it} = \epsilon_i + \epsilon_t + v_{it}$$

The determination of which model to use depends on two sets of tests.

Firstly, the appropriateness of the pooled regression method can be determined by testing for existence of heterogeneity using the Lagrange multiplier test by Breusch and Pagan (1980). Significant test statistics suggest unobserved heterogeneity and a pooled OLS regression would therefore be inappropriate because it lead to biased estimates (Sassen et al, 2016).

Secondly, once in the presence of evidence of heterogeneity, a determination between the fixed and random effects approaches can be made. The random effects approach is valid only when the composite error term ω_{it} is uncorrelated with all of the explanatory variables. The Hausman test can be utilized to determine whether this necessary assumption holds. If it does not, the random effects approach is inappropriate and the fixed effects approach should be followed (Brooks, 2014).

The determination of the most appropriate model, of those discussed above, to test each of the relationships studied in this thesis, along with accompanying diagnostic tests, is presented in the following model.

8.1.2 Diagnostic tests

It is necessary to perform diagnostic tests to ensure that the model chosen is appropriate and robust. Linear regression models are built upon assumptions about the data used. It is possible that the results from the model are unreliable if some of these assumptions are violated. This is could then lead to our inferences being unfounded.

Heteroscedasticity:

One of the assumptions in classical linear regression models is that the variance of the error terms is constant. This case is refered to homoscedasticity. The absence of homoscedasticity is referred to as heteroscedasticity (Brooks, 2014).

The problems created by heteroscedasticity are related to the OLS standard errors and regression inference. The OLS standard errors will likely be too large for the intercept. In the case that the variance of the error terms is positively related to the square of an explanatory variable, the OLS standard error for the slope will be low. In the case that the variance of the error terms is inversely related to the square of an explanatory variable, the OLS standard error for the slope will be large. These effects can lead to misleading inferences (Brooks, 2014).

There are a number of tests for heteroscedasticity. One such test is the Breusch–Pagan–Godfrey Test. The test assumes the error variance can be described as a linear function of nonstochastic Z variables, some of all of which can be the explanatory variables, such as follows (Gujarati & Porter, 2014):

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi}$$

If $\alpha_2 = \alpha_2 = \cdots = \alpha_m = 0$ then $\sigma_i^2 = \alpha_1$, which is a constant. A test of whether $\alpha_2 = \alpha_2 = \cdots = \alpha_m = 0$ is therefore a form a test to determine is σ_i^2 is homoscedastic.

The test is conducted by first obtaining the residuals from a linear regression and defining:

$$\tilde{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n}$$

$$p_i = \frac{\sum \hat{u}_i^2}{\tilde{\sigma}^2}$$

Then the constructures p_i are regressed up on the Z variables as follows:

$$p_i = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} + v_i$$

Then the explained sum of squares (ESS) from the regression above can be extracted and used to define the test statistic Θ as follows (Gujarati & Porter, 2014):

$$\Theta = \frac{1}{2}(\text{ESS})$$

$$\Theta \sim \chi_{m-1}^2$$

Autocorrelation:

Another assumption in classical linear regression models is that the covariance between error terms over time or cross-sectionally, depending on the model, is zero. The problem that arises when this is not the case is referred as autocorrelation or serial correlation of the error terms (Brooks, 2014).

The problems created by autocorrelation are related to the standard error estimates of the regression model. This could lead to incorrect regression inferences. In the presence of positive autocorrelation, the standard error estimates will be biased downwards, understating their true variability, and leading to an increase in probability of type I error. In addition, R^2will be inflated because autocorrelation leads to underestimation of the true error variance (Brooks, 2014).

Tests for autocorrelation focus on the residuals $\hat{u}$ in cases when the population disturbance terms are unobservable. The commonly used Durbin-Watson (DW) test is used to detect first order autocorrelation (i.e., the error term $\hat{u}_t$ and the error term $\hat{u}_{t-1}$immediately before it (Brooks, 194).

The DW test statistic can be define as follows (Brooks, 194):

$$DW = \frac{\sum_{t=2}^{T}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=2}^{T}\hat{u}_t}$$

The limits of the DW test statistic are 0 and 4. A lower limit of 0 corresponds with perfect positive autocorrelation in the residuals. The upper limit of 4 corresponds with perfect negative autocorrelation in the residuals. A DW test statistic in the middle of range at 2 corresponds with no autocorrelation in the residuals.

Determining statistical significance of any autocorrelation depends on two critical values between the limits of 0 and 4: a lower critical value $d_l$and upper critical value d_u. A DW test statistic in excess of either of these results in statistically significant autocorrelation, while an intermediate region before these critical values results in inconclusive evidence of autocorrelation in while the null hypothesis of no autocorrelation can be neither rejected nor not rejected (Brooks, 2014). These critical values are dependent on the number of observations and the number of, but not values taken, of the explanatory variables in the model (Gujarati & Porter, 2014).

A second and more general test for autocorrelation is the Breusch-Godfrey test. It allows examination of the relationship between $\hat{u}_t$ and many of its lagged values (i.e., $\hat{u}_{t-1}$, $\hat{u}_{t-2}$, and onwards) simultaneously. The model for the errors is as accordingly as follows:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \cdots + \rho_r u_{t-r} + v_t, \ \ v_t \sim N(0, \sigma_v^2)$$

The null and alternative hypothesis are:

$H_0 : \rho_1 = 0$ and $\rho_1 = 0$ and $\rho_1 = 0$ and …. and $\rho_r = 0$

$H_1 : \rho_1 \neq 0$ or $\rho_1 \neq 0$ or $\rho_1 \neq 0$ or …. or $\rho_r \neq 0$

The test is conducted by first obtaining the residuals from a linear regression and then secondly conducting the following regression:

$$\hat{u}_{t-1} = \gamma_1 + \gamma_2 x_{2t} + \gamma_3 x_{3t} + \gamma_4 x_{4t} + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \rho_3 \hat{u}_{t-3} + \cdots + \rho_r \hat{u}_{t-r} + v_t, \ \ v_t \sim N(0, \sigma_v^2)$$

Thirdly, the test statistic is given by

$$(T-r)R^2 \sim \chi_r^2$$

Where R^2 is the coefficient of determination above and T is the number of observations.

Multicollinearity:

Multicollinearity is a problem that occurs when the explanatory variables in a model are very highly correlated with each other. Two classes can be distinguished. Firstly, perfect multicollinearity is when an exact relationship exists between two or more explanatory variables, and is usually only observed when the same explanatory variable is erroneously included twice in a regression model. Near multicollinearity is a more common problem in practice. This occurs when there is a non-negligible but not perfect relationship between two or more explanatory variables (Brooks, 2014).

The primary issue introduced by multicollinearity is related to regression inference testing. The standard errors of the coefficients of the individual explanatory variables will be high due the difficulty of identifying the individual contribution of each explanatory variable to the fit of the regression model (Brooks, 2014).

Multicollinearity can be tested by examining the variance-inflating factors (VIF) of each explanatory variable in the regression model. In the case of a regression with k explanatory variables, the VIF for each independent variable would be defined as follows (Gujarati & Porter, 2014):

$$VIF_j = \frac{1}{1 - R_j^2}$$

When $R_i^2$is defined as the coefficient of determination $R^2$of a linear regression of that independent variable X_j on all of the remaining independent variables in the model. The VIF_j can therefore range from 1, which indicates no collinearity, to infinity as R_i^2 ranges from 0 (i.e., the case where no variance in the independent variable X_j is caused by the other independent variables) to 1 (i.e., the case where all of the variance in the independent variable X_j is caused by the other independent variables) (Gujarati & Porter, 2014).

Therefore, the higher the VIF_j the more troublesome and collinear X_j is considered to be. As a rule of thumb, a variable is considered is said to be highly collinear if its VIF_j is greater than 10, which corresponds to a R_j^2 of 0.90 (i.e., when greater than 90% of its variance is caused by the variance in the other independent variables) (Gujarati & Porter, 2014).

8.2 Model specification and hypothesis testing

8.2.1 CSP and idiosyncratic volatility

Firstly, the Lagrange multiplier test was applied on a pooled linear regression model of aggregate CSP and idiosyncratic volatility. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 21,117, 2,012, 23,129, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

A comparison of the fixed effects and random effects approaches was conducted using the Hausman test. The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate.

The Durbin-Watson (DW) test for first order autocorrelation was conducted. The test returned a highly statistically significant DW test statistic of 1.8807. In addition, the Breusch-Godfrey/Wooldridge test for higher order autocorrelation returned a highly statistically significant χ^2 test statistic of 40.023. Both tests indicate autocorrelation.

The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 5,924.3, which indicates the presence of heteroscedasticity.

The VIF of each explanatory variable in the model was examined in order to test for multicollinearity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the aggregate CSP and idiosyncratic volatility was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$idio_{it} = \beta_0 + \beta_1 ESG_{it} + \beta_2 Size_{it} + \beta_3 ROA_{it} + \beta_4 SDROA_{it} + \beta_5 Leverage_{it} + \beta_6 MTB_{it} + \beta_7 Liquidity_{it} + \beta_8 Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

Furthermore, in the presence of evidence suggesting both autocorrelation and heteroscedasticity in this model, the Newey-West procedure has been used to produce 'HAC' (heteroscedasticity and autocorrelation consistent) standard errors that correct for both the autocorrelation and heteroscedasticity that may be present (Brooks, 2014).

Research question 1a (i.e., is there a relationship between ESG measures and idiosyncratic risk?) is addressed using the following hypothesis test:

$$H_0\colon\ \beta_1 = 0$$

$$H_1:\ \beta_1 \neq 0$$

Secondly, the Lagrange multiplier test on the disaggregated CSP component and idiosyncratic volatility. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 21,022, 2,011, 23,033, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a highly statistically significant *DW* test statistic of 1.882. There was a significant χ^2 test statistic of 39.198. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 5,952.7, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

In conclusion, the relationship between the disaggregated components of CSP and idiosyncratic volatility was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$\begin{aligned} idio_{it} = \ & \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \beta_4 Size_{it} + \beta_5 ROA_{it} \\ & + \beta_6 SDROA_{it} + \beta_7 Leverage_{it} + \beta_8 MTB_{it} + \beta_9 Liquidity_{it} + \beta_{10} Div.pay_{i,t-1} + \mu_i \\ & + \lambda_t + v_{it} \end{aligned}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 1a is also addressed using the following hypothesis tests:

$$H_{01}:\ \gamma_1 = 0, H_{11}:\ \gamma_1 \neq 0$$

$$H_{02}:\ \gamma_1 = 0, H_{12}:\ \gamma_1 \neq 0$$

$$H_{03}:\ \gamma_1 = 0, H_{13}:\ \gamma_1 \neq 0$$

8.2.2 CSP and downside idiosyncratic volatility

Firstly, the Lagrange multiplier test was applied on aggregate CSP and downside idiosyncratic volatility. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 19,670, 1,759.8, and 21,430, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a highly statistically significant DW test statistic of 1.9053. There was a highly statistically significant χ^2 test statistic of 25.166. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 5,787.8, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the aggregate CSP and downside idiosyncratic volatility was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$ds.idio_{it} = \beta_0 + \beta_1 ESG_{it} + \beta_2 Size_{it} + \beta_3 ROA_{it} + \beta_4 SDROA_{it} + \beta_5 Leverage_{it} + \beta_6 MTB_{it} + \beta_7 Liquidity_{it} + \beta_8 Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 1b (i.e., is there a relationship between ESG measures and downside idiosyncratic risk?) is addressed using the following hypothesis test:

$$H_0\text{: } \beta_1 = 0$$

$$H_1\text{: } \beta_1 \neq 0$$

The Lagrange multiplier test was applied on disaggregated CSP component and downside idiosyncratic volatility. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 19,618, 1,759.2, and 21,377, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a highly statistically significant DW test statistic of 1.9066. There was a highly statistically significant χ^2 test statistic of 24.475. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 5,823.7, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the disaggregated components of CSP and downside idiosyncratic volatility was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$ds.idio_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \beta_4 Size_{it} + \beta_5 ROA_{it} + \beta_6 SDROA_{it} + \beta_7 Leverage_{it} + \beta_8 MTB_{it} + \beta_9 Liquidity_{it} + \beta_{10} Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 1b is also addressed using the following hypothesis tests:

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

8.2.3 CSP and upside idiosyncratic volatility

The Lagrange multiplier test was applied on aggregate CSP and upside idiosyncratic volatility. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 20,092, 1,996.8, and 22,089, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a highly statistically significant DW test statistic of 1.8808. There was a highly statistically significant χ^2 test statistic of 39.861. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 5,738.9, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the aggregate CSP and upside idiosyncratic volatility was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$up.idio_{it} = \beta_0 + \beta_1 ESG_{it} + \beta_2 Size_{it} + \beta_3 ROA_{it} + \beta_4 SDROA_{it} + \beta_5 Leverage_{it} + \beta_6 MTB_{it} + \beta_7 Liquidity_{it} + \beta_8 Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 1c (i.e., is there a relationship between ESG measures and upside idiosyncratic risk?) is addressed using the following hypothesis test:

$$H_0\colon \beta_1 = 0$$

$$H_1\colon\ \beta_1 \neq 0$$

The Lagrange multiplier test was on disaggregated CSP component and upside idiosyncratic volatility. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 19,959, 1,995.7, and 21,954, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a highly statistically significant DW test statistic of 1.8819. There was a highly statistically significant χ^2 test statistic of 39.159. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 5,759.6, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the disaggregated components of CSP and upside idiosyncratic volatility was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$\begin{aligned} up.idio_{it} = {} & \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \beta_4 Size_{it} + \beta_5 ROA_{it} \\ & + \beta_6 SDROA_{it} + \beta_7 Leverage_{it} + \beta_8 MTB_{it} + \beta_9 Liquidity_{it} + \beta_{10} Div.pay_{i,t-1} + \mu_i \\ & + \lambda_t + v_{it} \end{aligned}$$

‘HAC’ standard errors were used (Brooks, 2014).

Research question 1c is also addressed using the following hypothesis tests:

$$H_{01}\colon\ \gamma_1 = 0, H_{11}\colon\ \gamma_1 \neq 0$$

$$H_{02}\colon\ \gamma_1 = 0, H_{12}\colon\ \gamma_1 \neq 0$$

$$H_{03}\colon\ \gamma_1 = 0, H_{13}\colon\ \gamma_1 \neq 0$$

Finally, research question 1d (i.e., How do these relationships, or lack thereof, compare with each other?) is addressed by examining the results of the preceding hypothesis tests and the strengths and directions of the associations suggested by the regression coefficients.

8.2.4 CSP and idiosyncratic skewness

The Lagrange multiplier test was applied on aggregate CSP and idiosyncratic skewness. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 340.8, 0.91554, and 341.71, respectively. The test statistics for time and two ways effects are highly statistical significant. The test statistic for firm effects is not statistically significant. This suggests time and twoway effects. To confirm, an F test for time effects was conducted which returned a highly statistically significant F statistic of 8.2724. An F test for twoway effects returned a statistically insignificant F statistic of 1.0481. This suggests time effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic not statistically significant. The fixed effects approach is therefore more appropriate. There was a highly statistically significant DW test statistic of 1.9445. There was a highly statistically significant χ^2 test statistic of 8.3195. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a statistically insignificant test statistic of 0.485, which suggests the lack of presence of heteroscedasticity.

The relationship between the aggregate CSP and idiosyncratic skewness was estimated using the following linear regression model with time-fixed effects:

$$is_{it} = \beta_0 + \beta_1 ESG_{it} + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Utilizing the connection between idiosyncratic skewness and the proportion of each component of idiosyncratic risk to total idiosyncratic risk, research question 1d will also be partially addressed using the following hypothesis test:

$$H_0\colon \beta_1 = 0$$

$$H_1\colon \beta_1 \neq 0$$

The Lagrange multiplier test was on disaggregated CSP component and idiosyncratic skewness. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 332.2, 0.8414, and 333.05, respectively. The test statistics for time and two ways effects are highly statistical significant. The test statistic for firm effects is not statistically significant. This suggests time and twoway effects. To confirm, an F test for time effects was conducted which returned a highly statistically significant F statistic of 8.1922. An F test for twoway effects returned a statistically insignificant F statistic of 1.0507. This suggests time effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was not statistically significant. The random effects approach is therefore more appropriate. There was a highly statistically significant DW test statistic of 1.9435. There was a highly statistically significant χ^2 test statistic of 8.6428. Both tests indicate autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a statistically insignificant test statistic of 3.874, which suggests the lack of presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the disaggregated components of CSP and idiosyncratic skewness was estimated using the following multiple linear regression model with random time effects:

$$is_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \omega_{it}, \quad \omega_{it} = \epsilon_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 1c is also addressed using the following hypothesis tests:

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

8.2.5 CSP and abnormal returns

The Lagrange multiplier test was on aggregate CSP and abnormal returns. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 2,838, 3.8752, and 2,841.8, respectively. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a statistically insignificant *DW* test statistic of 2.0897. In contrast, the Breusch-Godfrey/Wooldridge test for higher order autocorrelation returned a highly statistically significant χ^2 test statistic of 21.8. This suggests autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 196.2, which indicates the presence of heteroscedasticity.

The relationship between the aggregate CSP and abnormal returns was estimated using the following linear regression model with firm-fixed and time-fixed effects:

$$\alpha_{it} = \beta_0 + \beta_1 ESG_{it} + \mu_i + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 2a (i.e., there a relationship between ESG measures and market return?) is addressed using the following hypothesis test:

$$H_0\colon\ \beta_1 = 0$$

$$H_1\colon\ \beta_1 \neq 0$$

The Lagrange multiplier test was applied on disaggregated CSP component and abnormal returns. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 2,852.3, 3.9329, and 2,856.2, respectively. The test statistics for time and two ways effects are highly statistical significant. All three test statistics are highly statistical significant and suggest both individual and firm effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a statistically insignificant *DW* test statistic of 2.0897. There was a highly statistically significant χ^2 test statistic of 21.808. This suggests autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 254.6, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the disaggregated components of CSP and abnormal returns was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$\alpha_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \mu_i + \lambda_t + v_{it}$$

Research question 2a is also addressed using the following hypothesis tests:

$$H_{01}\colon\ \gamma_1 = 0, H_{11}\colon\ \gamma_1 \neq 0$$

$$H_{02}\colon\ \gamma_1 = 0, H_{12}\colon\ \gamma_1 \neq 0$$

$$H_{03}\colon\ \gamma_1 = 0, H_{13}\colon\ \gamma_1 \neq 0$$

'HAC' standard errors were used (Brooks, 2014).

8.2.6 CSP and excess log returns

The Lagrange multiplier test was applied on aggregate CSP component and excess log returns. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 390,610, 0.17923, 390,610, respectively. The test statistics for time and two ways effects are highly statistical significant. The test statistic for firm effects is not statistically significant. This suggests time and twoways effects. To confirm, an F test for time effects was conducted which returned a highly statistically significant F statistic of 415.33. An F test for twoway effects was conducted which returned a highly statistically significant F statistic of 6.74. This suggests twoway effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was statistically insignificant *DW* test statistic of 2.0774. In contrast, the Breusch-Godfrey/Wooldridge test for higher order autocorrelation returned a highly statistically significant χ^2 test statistic of 16.235. This suggests autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 159.4, which indicates the presence of heteroscedasticity.

The relationship between the aggregate CSP and excess log returns was estimated using the following linear regression model with firm-fixed and time-fixed effects:

$$r_{it} = \beta_0 + \beta_1 ESG_{it} + \mu_i + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 2a is also addressed using the following hypothesis test:

$$H_0\colon\ \beta_1 = 0$$

$$H_1\colon\ \beta_1 \neq 0$$

The Lagrange multiplier test was applied on disaggregated CSP components and excess log returns. The results for time effects, firm effects, and two way effects returned χ^2 test statistics of 389,600, 0.13591, and 389,600, respectively. The test statistics for time and two ways effects are highly statistical significant. The test statistic for firm effects is not statistically significant. This suggests time and twoway effects. To confirm, an F test for time effects was conducted which returned a highly statistically significant F statistic of 359.9. An F test for twoway effects was conducted which returned a highly statistically significant F statistic of 6.734. This suggests twoway effects. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a statistically insignificant *DW* test statistic of 2.0777. In contrast, the Breusch-Godfrey/Wooldridge test for higher order autocorrelation returned a highly statistically significant χ^2 test statistic of 16.377. This suggests autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 213.3, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

The relationship between the disaggregated components of CSP and excess log returns was estimated using the following multiple linear regression model with firm-fixed and time-fixed effects:

$$r_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \mu_i + \lambda_t + v_{it}$$

'HAC' standard errors were used (Brooks, 2014).

Research question 2a is also addressed using the following hypothesis tests:

$$H_{01}: \gamma_1 = 0, H_{11}: \gamma_1 \neq 0$$

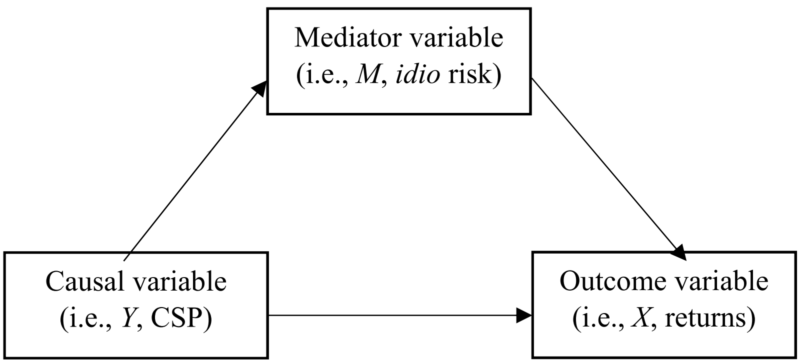
$$H_{02}: \gamma_1 = 0, H_{12}: \gamma_1 \neq 0$$

$$H_{03}: \gamma_1 = 0, H_{13}: \gamma_1 \neq 0$$

8.3 Mediation analysis

Finally, a mediation analysis can be conducted (Kenny, n.d.) following the methodology developed by Baron & Kenny (1986) to determine whether any relationship found between CSP and returns is mediated by *idio* risk. This can address research question 2c.

Adopting their terminology, the three variables in the model are:



Establishing mediation consists of four steps:

1. It must be established that there is an effect to may be mediated. This requires regressing Y on X. In our case, regressing returns on CSP.
2. It must be established that the causal variable is correlated with the mediator variable. This requires regressing M on X. In our case, regressing *idio* on CSP.

3. It must be established that the mediator variable affects the outcome variable. This requires regressing Y simultaneously X and M. In our case, conducting a regression analysis of returns using both CSP and *idio* as explanatory variables.
4. To establish total mediation, the effect of X on Y controlling for M should be zero. In our case, total mediation would be demonstrated if the coefficient of CSP in the regression on returns becomes statistically insignificant when *idio* is added to the regression model.

There is evidence of partial mediation in the data if steps one through three are achieved but not step four. In this case, the coefficient estimated for CSP in the regression model when *idio* is added is remain statistically significant but reduce in strength. This would constitute evidence that *Idio* partially explains (i.e., mediates) some of the effect of CSP on returns.

9 Empirical results and discussion

9.1 Descriptive statistics

Table 1 contains descriptive statistics of the risk measures, return measures, ESG measures, and control variables.

In panel A presenting risk measures, mean *idio* is 0.247 with a standard deviation of 0.161. This is comparable to Sassen et al. (2016) at 0.33 and standard deviation of 0.1734. A lower *idio* would be consistent with greater predictive power of the French Fama five factor model used in this thesis in comparison to the Carhart four factor used in Sassen et al. (2016). However, this difference in predictive power is not conclusive and there are other methodological differences that could also contribute to the lower mean *idio* found in this sample (e.g., an updated longer time period that adds 2013 through 2016).

Table 1: Descriptive statistics

Variable	N	Mean	St. Dev.	Min	Max	Skewness	Kurtosis
Panel A: risk measures							
idio	10,827	0.247	0.161	0.022	2.814	3.447	27.013
ds.idio	10,827	0.173	0.115	0.014	2.254	3.613	31.269
up.idio	10,827	0.174	0.116	0.017	1.696	3.337	24.227
is	10,827	0.008	0.245	-1.086	1.030	0.031	3.5860
Panel B: return measures							
r	10,827	0.015	0.439	-5.791	2.832	-1.63	13.225
alpha	10,827	-0.001	0.053	-0.759	0.630	-0.951	20.709
Panel C: ESG measures							
ESG	10,827	0.619	0.239	0.045	0.971	-0.633	2.2980
SOCSCORE	10,827	0.657	0.282	0.034	0.990	-0.688	2.1630
CGVSCORE	10,827	0.556	0.277	0.012	0.978	-0.339	1.8760
ENVSCORE	10,827	0.643	0.291	0.085	0.975	-0.598	1.8570
Panel C: control variables							
Size	10,827	15.834	1.887	7.412	22.052	0.543	3.2340
ROA	10,827	0.067	0.130	-2.460	3.158	4.396	173.188
SDROA	10,827	0.044	0.066	0.0002	1.417	7.090	94.013
Leverage	10,827	0.192	0.170	0.000	2.515	2.369	20.956
MTB	10,827	2.658	2.873	-3.027	19.013	3.023	15.686
Liquidity	10,827	0.906	0.847	0.002	4.717	1.938	7.9360
Div.pay.1	10,827	0.023	0.051	0.000	0.435	6.150	47.266

Idio has positive skewness and excess kurtosis, indicating that firm-specific shocks are more commonly positive than negative in our sample, and excess kurtosis, indicating that extreme firm-specific shocks are more common that would be expected if they followed a normal distribution. This results are in line Sassen et al (2016).

Ds.idio and *up.idio* are very similar. *Ds.idio* has a of 0.173 and a standard deviation of 0.115, which is only slightly lower than the case of *up.idio* with a mean of 0.174 and a standard deviation of 0.116. is has a mean of 0.008 and a standard deviation of 0.245.

In panel B presenting return measures, *r* has a mean of a 0.015 and a standard deviation of 0.439, indicating a relatively large amount of variance. Overall, *r* has a negative skew, indicating greater likelihood of negative returns in the distribution. This is consistent with findings that historical returns on stocks exhibit more frequent large negative deviations from the mean than would be expected if truly normally distributed (Bodie, 2014)

Furthermore, *r* has a kurtosis far greater than the kurtosis of three exhibited by the normal distribution, indicating a higher likelihood in our sample of extreme returns at either end of the distribution. This is consistent with the findings that stock return distributions in general have heavy tails: very low and very high returns are observed more frequently than they would be if they followed a normal distribution (Munk, 2016).

These insights are reassuring because there would be no point of decomposing risk measures related to dispersion around central tendencies into their downside and upside components if the returns exhibited a normal distribution (i.e., zero skewness). The distribution would be symmetrical in this case and *ds.idio* and *up.idio* would be the same value.

The descriptive statistics of alpha are similar to those of *r* but lower in magnitude (e.g., mean almost at zero, negative skewness) but with higher kurtosis, which is generally consistent with the interpretation of alpha as the abnormal returns that remain of *r* after controlling for returns attributable to exposure to common risk factors.

In panel C presenting ESG measures, the three components *SOCSCORE*, *CGVSCORE*, and *ENVSCORE* share similar means of 0.657, 0.556, and 0.643, respectively, and similar standard deviations of 0.282, 0.277, and 0.291, respectively. The descriptive statistics of ESG are similar due to its construction as an equally weighted sum of the components. This statistics are very similar to those of Sassen et al (2016), suggesting that Asset4 ESG measures remain stable over time. This may be because socially responsible firms do not rapidly change their level of social responsibility, or due to the specific methodology of Asset4 ESG measures, or some combination of those two factors with others.

In panel C we see that the descriptive statistics of the control variances, which are similar to those of Sassen et al (2016). These descriptive statistics were calculated after the process of winsorization at the 1% level was conducted on MTB, Liquidity, and Div.pay.1, which contains large outliers, as described in the data section above.

9.2 Correlation table

Table 2 contains the Person product-moment correlation coefficients and accompanying statistical significance test results of the risk measures, return measures, ESG measures, and control variables. I adopt the terminology of some authors of statistical textbooks that suggest that correlations coefficients between -0.20 and +0.20 indicate a weak association between two variables, those between 0.20 and 0.50 in absolute terms indicate a moderate association, and those larger than 0.50 in absolute terms indicate a strong association (Urdan, 2010).

Focusing first on the primary relationship of study, CSP and risk measures, there is a highly statistically significant weak negative correlation between *idio* and ESG. The same is found for *idio* and the ESG components. This is consistent with theories predicting that CSP lowers firm risk. Decomposing *idio* into *ds.idio* and *up.idio* does not yield significantly different results. In each case, the direction and statistical significance of the correlation coefficient is the same. The strength of the correlation coefficient is substantially the same in each case as well. This lack of distinguish may be evident in the lack of statistically significant correlation between *is* and any of the risk measures.

Focusing next on the other primary relationship of study, CSP and returns, there are no statistically significant correlation coefficients shared by any combination of ESG measure and risk measure. This suggests a lack of association between CSP and returns.

Moving onto the other relationships, it is also encouraging to see that the risk measures share a highly statistically significant correlation with all of the control variables, suggesting that control variables are justified additions to the linear regressions to follow.

The strength of the correlation coefficient between *ENVSCORE* and *SOCSCORE* is a highly statistically significant 0.7791, which indicates a strong association and may suggest multicollinearity. This motivates the test for multicollinearity performance via analysis of variance inflation factors.

Table 2: Correlation table

	idio	*ds.idio*	*up.idio*	is	*r*	alpha	ESG	*SOCSCORE*	*CGVSCORE*	*ENVSCORE*	Size	ROA	SDROA	Leverage	MTB	Liquidity	Div.pay.1
idio	1.0000																
ds.idio	0.9849***	1.0000															
up.idio	0.9851***	0.9406***	1.0000														
is	0.0070	-0.1351***	0.1477***	1.0000													
r	-0.3419***	-0.3354***	-0.3383***	-0.0175*	1.0000												
alpha	-0.2171***	-0.2175***	-0.2101***	-0.0048	0.4385***	1.0000											
ESG	-0.1128***	-0.1091***	-0.1130***	-0.0108	-0.0014	-0.0075	1.0000										
SOCSCORE	-0.1206***	-0.1183***	-0.1190***	-0.0062	-0.0062	-0.0045	0.8963***	1.0000									
CGVSCORE	-0.0481***	-0.0440***	-0.0507***	-0.0155	0.0113	-0.0111	0.7493***	0.4802***	1.0000								
ENVSCORE	-0.1150***	-0.1119***	-0.1144***	-0.0058	-0.0083	-0.0034	0.8782***	0.7791***	0.4258***	1.0000							
Size	-0.1474***	-0.1453***	-0.1444***	0.0060	-0.0713***	-0.0163*	0.3490***	0.3767***	0.1234***	0.3761***	1.0000						
ROA	-0.1849***	-0.1815***	-0.1820***	0.0007	0.1648***	0.1222***	-0.0286***	-0.0274***	0.0027	-0.0464***	-0.1836***	1.0000					
SDROA	0.2638***	0.2601***	0.2597***	-0.0028	-0.0663***	-0.0462***	-0.1307***	-0.1670***	-0.0065	-0.1536***	-0.2836***	0.0331***	1.0000				
Leverage	0.0611***	0.0610***	0.0591***	0.0029	-0.0628***	-0.0554***	0.0618***	0.0425***	0.0740***	0.0403***	-0.0064	-0.1140***	-0.0122	1.0000			
MTB	-0.0987***	-0.0990***	-0.0958***	0.0034	0.1750***	0.1132***	-0.0050	0.0000	0.0464***	-0.0563***	-0.2345***	0.3821***	0.0788***	-0.0605***	1.0000		
Liquidity	0.1379***	0.1361***	0.1364***	-0.0059	-0.0557***	-0.0297***	0.1761***	0.1317***	0.2305***	0.0861***	0.0257***	-0.0520***	0.1028***	0.0995***	-0.0073	1.0000	
Div.pay.1	-0.0305***	-0.0336***	-0.0263***	0.0121	-0.0669***	-0.0011	-0.0144	0.0443***	-0.1111***	0.0275***	0.1730***	-0.0212**	-0.0842***	0.0820***	-0.0812***	-0.0349***	1.0000

Furthermore, results show a highly statistically significant strong correlation coefficient of 0.9606 between *ds.idio* and *up.idio*, a highly statistically significant strong correlation coefficient of 0.9849 between *idio* and *ds.idio*, and highly statistically significant strong correlation coefficient of 0.9851 between *idio* and *up.idio*. This suggests that these measures may be measuring the same phenomena.

This is consistent with the findings of Koch (2010) using a sample of German listed firms during the period of 1974 to 2016 and residuals from the CAPM. He found that average cross-sectional correlation between idiosyncratic volatility and its upside and downside measures were 0.948 and 0.905 over time, respectively, and 0.725 between downside and upside idiosyncratic volatility over time. This led him to conclude that concludes that *ds.idio* and *up.idio* capture essentially the same risk.

This is stark contrast to the findings of Frieder & Jiang (2007) using a sample of American listed stocks (i.e., listed on the NYSE, AMEX, and NASDAQ) during the period 1980 to 2005 and residuals from the Carhart four-factor model. They found that the mean and median number cross-sectional correlations between up and down idiosyncratic volatility at the monthly frequency was 0.19 and 0.18. The same mean and median correlations at the quarterly frequency were 0.14 and 0.12, respectively.

It is not clear what the cause of the great discrepancy is but it does not appear related to the methodology of calculating correlation coefficients. Applying the methodology of Koch (2010) (i.e., first taking the cross-sectional correlations of *idio*, *ds.idio*, and *up.idio* each year and then taking the mean of these yearly correlations coefficients) to my sample results in a *idio* to *ds.idio* correlation coefficient of 0.9797, *idio* to *up.idio* correlation coefficient of 0.9813, and *ds.idio* to *up.idio* correlation coefficient of 0.9232.

Following the methodology of Frieder & Jiang (2007) (i.e., taking the cross sectional mean of *idio*, *ds.idio*, and *up.idio* each year and then computing the correlation of these means to each other) to my sample results in an *idio* to *ds.idio* correlation coefficient of 0.9981, *idio* to *up.idio* correlation coefficient of 0.9983, and *ds.idio* to *up.idio* correlation coefficient of 0.9929.

The correlation coefficients between the isio, *ds.idio*, and *up.idio* risk measures and return measures (i.e., *r* and alpha) are negative and highly statistically significant. This does not support the theories that found that *idio* carries a risk premium. However, correlation coefficients are only a necessary but not sufficient condition of determining a predictive relationship.

Correlations are insightful but only measure association. They do not distinguish between explanatory and dependent variables. These measures do not take into account predictive power (i.e., the direction of the relationship) (Urdan, 2010). For an analysis of prediction, we must move onto the closely related family of regression analysis.

9.3 Regressions results and discussion

Table 3 contains the results of the regressions conducted in order to estimate the predictive power of the explanatory variables of interest (i.e., ESG measures) on the dependent variables studied (i.e., risk measures, is, and return measures).

9.3.1 CSP and idiosyncratic volatility

Firstly, a multiple linear regression with firm-fixed and time-fixed effects was conducted to examine the predictors of idiosyncratic volatility $idio$. Eight explanatory variables were simultaneously entered in the model: our variable of interest, *ESG*, and seven control variables to help isolate the predictive power of *ESG*.

$$idio_{it} = \beta_0 + \beta_1 ESG_{it} + \beta_2 Size_{it} + \beta_3 ROA_{it} + \beta_4 SDROA_{it} + \beta_5 Leverage_{it} + \beta_6 MTB_{it} + \beta_7 Liquidity_{it} + \beta_8 Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

$$\mathrm{H}_0\text{: } \beta_1 = 0$$

$$\mathrm{H}_1\text{: } \beta_1 \neq 0$$

In total, the model accounted for 46.99% of the variance in *idio* and the F statistic indicates that model is highly statistically significant overall.

I fail to reject H_0. β_1 is not statistically significantly different from zero. There is no statistically significant evidence to suggest that the relationship between ESG and $idio$ is different from zero (i.e., that none exists). *ESG* was not a statistically significant predictor in the model.

Of the control variables, *ROA*, *SDROA*, *Leverage*, *MTB*, and *Div.pay.1* were statistically significant predictors of *idio*. *ROA*, *MTB*, and *Div.pay.1* were negatively associated while *SDROA* and *Leverage* were positively associated. Size was a marginally statistically significant negative predictor. Liquidity was not a statistically significant predictor.

Table 3: Regression analyses

	Dependent variable:											
	idio	idio	ds.idio	ds.idio	up.idio	up.idio	iv	iv	α	α	*r*	*r*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
SOCSCORE	-0.0390***		-0.0286***		-0.0262**		-0.0003		-0.0040		-0.0155	
	(0.0150)		(0.0106)		(0.0108)		(0.0137)		(0.0054)		(0.0362)	
CGVSCORE	0.0097		0.0106		0.0036		-0.0066		-0.0083*		-0.0442	
	(0.0126)		(0.0091)		(0.0090)		(0.0098)		(0.0046)		(0.0319)	
ENVSCORE	0.0146		0.0114		0.0091		0.0018		-0.0074		-0.0552*	
	(0.0138)		(0.0096)		(0.0100)		(0.0130)		(0.0053)		(0.0334)	
ESG		-0.0158		-0.0077		-0.0140		-0.0026		-0.0195***		-0.1148***
		(0.0145)		(0.0102)		(0.0105)		(0.0099)		(0.0055)		(0.0398)
Size	-0.0096*	-0.0095*	-0.0059	-0.0058	-0.0076*	-0.0075*						
	(0.0056)	(0.0056)	(0.0040)	(0.0040)	(0.0040)	(0.0041)						
ROA	-0.1805***	-0.1809***	-0.1226***	-0.1229***	-0.1308***	-0.1310***						
	(0.0676)	(0.0676)	(0.0457)	(0.0458)	(0.0499)	(0.0498)						
SDROA	0.3516***	0.3527***	0.2440***	0.2448***	0.2529***	0.2537***						
	(0.0463)	(0.0466)	(0.0322)	(0.0325)	(0.0340)	(0.0341)						
Leverage	0.0851**	0.0844**	0.0628**	0.0622**	0.0562**	0.0558**						
	(0.0362)	(0.0364)	(0.0248)	(0.0249)	(0.0268)	(0.0269)						
MTB	-0.0035***	-0.0035***	-0.0027***	-0.0027***	-0.0023***	-0.0023***						
	(0.0011)	(0.0011)	(0.0008)	(0.0008)	(0.0008)	(0.0008)						
Liquidity	0.0031	0.0032	0.0016	0.0018	0.0027	0.0028						
	(0.0039)	(0.0039)	(0.0028)	(0.0028)	(0.0028)	(0.0028)						
Div.pay.1	-0.5778***	-0.5802***	-0.4255***	-0.4273***	-0.3893***	-0.3908***						
	(0.0947)	(0.0947)	(0.0651)	(0.0652)	(0.0700)	(0.0700)						
R^2 (total)	0.4705	0.4699	0.4569	0.4562	0.4627	0.4623		0.0115	0.1591	0.1591	0.4256	0.4256
F	47.62***	51.30***	47.04***	50.60***	45.69***	49.31***		8.10***	24.68***	27.67***	144.5***	161.2***

Note: $^{*}p<0.1$ $^{**}p<0.05$ $^{***}p<0.01$

These results are generally consistent with the relationships expected. An increase in operating profitability *ROA* is related to a decrease in *idio*, which is consistent with more profitable firms being more resilient to firm specific shocks. An increase in the volatility of operation profitability is related to an increase in *idio*, which is consistent with these two measures of risk being related. An increase in *Leverage* is related to an increase in *idio*, which in consistent with leverage increasing the inherent riskiness of a firm's capital structure and resulting cash flows to equity holders. An increase in *MTB* is related to a decrease in *idio*. This relation can be interpreted as the effect of *idio* increasing as the ratio of a firm's market value to its book value increases (i.e., as it tilts more towards a growth stock as opposed to a value stock). An increase in *Div.pay.1* is related to a decrease in *idio*, which is consistent with dividends being used as signals for managements inside knowledge of the strength and stability of future firm cash flows.

Not shown in the table to conserve space, the time-fixed effects on the 2008 and 2009 years were both highly statistically significant with a positive association of 0.09238 and 0.05727, respectively. The fixed effects on 2004 through 2007, 2010, 2013, 2014, and 2015 were statistically significant with a negative association. The fixed effects on 2011, 2012, and 2016 were marginally statistically significant with a negative association. The fixed effects on 2002 and 2003 were not statistically significant.

These results indicate that *idio* rose in 2008 and 2009 cross sectionally across our sample. This makes intuitive sense as these were years of high volatility in general due to the effects of the financial crisis.

These results are not consistent with Sassen et al (2016). Their results indicate that their analogous model explained 35.8% of idiosyncratic volatility. These are relative similar R^2 coefficients that may be explained by the differing asset pricing models used (i.e., Fama French five factor in comparison to the Carhart four-factor model in their case) and the additional years of data used in my model. However, their study found aggregate ESG to be highly statistically significant with a negative association of -0.041. In addition, only Size of the control variables was found to be a statistically significant predictor at -0.016, while ROA was a marginally positively associated and statistically significant 0.066.

Aggregating all components of CSP may obscure relationships between its components, which may be empirically and conceptually distinct concepts. It may also be the reason for the lack of statistically significant coefficient for *ESG* if the relations between the components and *idio* interact in some way as to cancel eachother other out when viewed in aggregate.

Accordingly, a multiple linear regression with firm-fixed and time-fixed effects was conducted to examine the predictors of *idio* using ten explanatory variables: our variables of interest *SOCSCORE*, *CGVSCORE*, and *ENVSCORE* and the same seven control variables.

$$idio_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \beta_4 Size_{it} + \beta_5 ROA_{it} + \beta_6 SDROA_{it} + \beta_7 Leverage_{it} + \beta_8 MTB_{it} + \beta_9 Liquidity_{it} + \beta_{10} Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

In total, the model accounted for 47.05% of the variance in *idio* and was highly statistically significant.

I fail to reject H_{02} and H_{03}. There is no statistically significant evidence to suggest that the relationship between $CGVSCORE$ and $idio$ or $ENVSCORE$ and $idio$ is different from zero. Neither *CGVSCORE* nor *ENVSCORE* were statistically significant predictors in the model.

I reject H_{01} in favour of H_{11}. There is statistically significant evidence to suggest that a relationship exists between $SOCSCORE$ and $idio$. *SOCSCORE* was a negatively associated highly statistically significant predictor ($\beta_1 = -0.0390$) in the mode.

These results indicate that, holding all other explanatory variables constant and controlling for time and firm effects, an increase in 0.01 in yearly *SOCSCORE* (i.e., 1 point in the raw 0 to 100 Asset4 scale) predicts a decrease of annualized idiosyncratic risk *idio* of 0.039%. Equivalently, an increase in one standard deviation of *SOCSCORE* (i.e., 0.282 or 28.2 in the original Asset4 scale) predicts a decrease in *idio* of 1.1%. This is in the context of average *idio* during this time period of 24.7%.

Theoretically, these results are as expected and consistent with theories conceptualizing the effect of CSP on firms as risk reducing via risk mitigation, an insurance-effect, or reduced news hoarding because of increased corporate governance.

While it is unfounded to make inferences from statistically insignificant results, it is interesting to see that the predictive effects of *CGVSCORE* and *ENVSCORE* run in opposite direction to *SOCSCORE*. If it were the case that *SOCSCORE* decreases *idio* while *CGVSCORE* and *ENVSCORE* increase it, this may contribute to the lack of predictive power of the aggregate ESG. This may be a case of countervailing effects that effectively cancel each other out, or reducing the negative association of the aggregate ESG to such a small degree as to no longer be statistically significant.

Of the control variables, the relations were similar to those at the aggregate ESG level. ROA, SDROA, Leverage, MTB, and Div.pay.1 were statistically significant predictors of *idio*. ROA, MTB, and

Div.pay.1 were negatively associated while SDROA, and Leverage were positively associated. Size was a marginally statistically significant negative predictor. Liquidity was not a statistically significant predictor. The coefficients were of the same statistical significance direction of association as in the case of *idio* regressed on the aggregate ESG and of very similar strength.

The time-fixed effects were the same statistical significance and direction as those described in the model using the aggregate ESG measure. The strength of the fixed-effect in each year were very similar.

These results are generally consistent with Sassen et al (2016), which also found *SOCSCORE* to be a statistically significant and negative predictor ($\beta = -0.021$). While this suggests the same type of relationship, the strength is a full 85.7% larger in our analysis. *ENVSCORE* is marginally statistically significant in their analysis and of opposite direction.

<u>*Sub-conclusion, research sub-question 1a*</u>

In conclusion, the results of the tests of these hypotheses address research question 1a. There is a significant relationship between *SOCSCORE* and idiosyncratic risk but no such relationship between either aggregate ESG, *CGVSCORE*, or *ENVSCORE* and idiosyncratic risk.

9.3.2 CSP and downside idiosyncratic volatility

Secondly, *idio* was decomposed into its downside component *ds.idio* and upside component *up.idio* in order to examine whether the explanatory variables ESG measures had differing relations with each.

A multiple linear regression with firm-fixed and time-fixed effects was conducted of *ds.idio* on ESG and the seven control variables.

$$ds.idio_{it} = \beta_0 + \beta_1 ESG_{it} + \beta_2 Size_{it} + \beta_3 ROA_{it} + \beta_4 SDROA_{it} + \beta_5 Leverage_{it} + \beta_6 MTB_{it} + \beta_7 Liquidity_{it} + \beta_8 Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

$$H_0\colon\ \beta_1 = 0$$

$$H_1\colon\ \beta_1 \neq 0$$

In total, the model accounted for 45.62% of the variance in *ds.idio* and was highly statistically significant.

I fail to reject H_0. β_1 is not statistically significantly different from zero. ESG was not a statistically significant predictor in the model.

Of the control variables, ROA, SDROA, Leverage, MTB, and Div.pay.1 were statistically significant predictors of *ds.idio*. ROA, MTB, and Div.pay.1 were negatively associated while SDROA and Leverage were positively associated. Size and Liquidity were not statistically significant predictors. The coefficients were of the same statistical significance direction of association as in the case of *idio*, but of a weaker strength in each case.

The time-fixed effects on the 2008 and 2009 year were both highly statistically significant with a positive association of 0.06401 and 0.04577, respectively. The fixed effects on 2004 through 2007, 2010, 2013, 2014, and 2015 were statistically significant with a negative association. The fixed effects on 2015 were marginally statistically significant with a negative association. The fixed effects on 2002, 2003, 2011, 2012, and 2016 were not statistically significant.

A multiple linear regression with firm-fixed and time-fixed effects was conducted of *ds.idio* on *SOCSCORE*, *CGVSCORE*, *ENVSCORE* and the seven control variables.

$$ds.idio_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \beta_4 Size_{it} + \beta_5 ROA_{it} + \beta_6 SDROA_{it} + \beta_7 Leverage_{it} + \beta_8 MTB_{it} + \beta_9 Liquidity_{it} + \beta_{10} Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

In total, the model accounted for 45.62% of the variance in *ds.idio* and was highly statistically significant.

I fail to reject H_{02} and H_{03}. There is no statistically significant evidence to suggest that the relationship between $CGVSCORE$ and $idio$ or $ENVSCORE$ and $idio$ is different from zero. Neither *CGVSCORE* nor *ENVSCORE* were statistically significant predictors.

I reject H_{01} in favour of H_{11}. There is statistically significant evidence to suggest that a relationship exists between $SOCSCORE$ and $idio$. *SOCSCORE* was a negatively associated highly statistically significant predictor ($\beta = -0.0286$).

Theoretically, these results are as expected and consistent with theories conceptualizing the effect of CSP on firms as risk reducing via risk mitigation, an insurance-effect, or reduced news hoarding because of increased corporate governance.

Of the control variables, ROA, SDROA, Leverage, MTB, and Div.pay.1 were statistically significant predictors of *ds.idio*. ROA, MTB, and Div.pay.1 were negatively associated while SDROA and Leverage were positively associated. Size and Liquidity were not statistically significant predictors. The coefficients were of the same statistical significance direction of association as in the case of *ds.idio* regressed on the aggregate ESG and of very similar strength.

The time-fixed effects were the same statistical significance and direction as those described in the model using the aggregate ESG measure. The strength of the fixed-effect in each year were very similar.

Sub-conclusion, research sub-question 1b

In conclusion, the results of the tests of these hypotheses address research question 1b. There is significant relationship between *SOCSCORE* and downside idiosyncratic risk but no such relationship between either aggregate ESG, *CGVSCORE*, or *ENVSCORE* and downside idiosyncratic risk.

9.3.3 CSP and upside idiosyncratic volatility

A multiple linear regression with firm-fixed and time-fixed effects was conducted of *up.idio* on ESG and the seven control variables.

$$up.idio_{it} = \beta_0 + \beta_1 ESG_{it} + \beta_2 Size_{it} + \beta_3 ROA_{it} + \beta_4 SDROA_{it} + \beta_5 Leverage_{it} + \beta_6 MTB_{it} + \beta_7 Liquidity_{it} + \beta_8 Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

$$H_0\colon\ \beta_1 = 0$$

$$H_1\colon\ \beta_1 \neq 0$$

In total, the model accounted for 46.27% of the variance in *up.idio* and was highly statistically significant.

I fail to reject H_0. β_1 is not statistically significantly different from zero. ESG was not a statistically significant predictor.

Of the control variables, ROA, SDROA, Leverage, MTB, and Div.pay.1 were statistically significant predictors of *up.idio*. ROA, MTB, and Div.pay.1 were negatively associated while SDROA and Leverage were positively associated. Size and Liquidity were not statistically significant predictors. The coefficients were of the same statistical significance direction of association as in the case of *idio*, but of a weaker strength in each case. In comparison to *ds.idio*, the coefficients were of the same statistical significance direction of association, and of generally similar strength.

The time-fixed effects on the 2008 and 2009 years were both highly statistically significant with a positive association of 0.06654 and 0.03538, respectively. The fixed effects on 2004 through 2007, 2010, 2011, 2012, 2013, 2014, and 2015 were statistically significant with a negative association. The fixed effects on 2002 and 2016 was marginally statistically significant with a positive and a negative association, respectively. The fixed effects on 2003 were not statistically significant.

A multiple linear regression with firm-fixed and time-fixed effects was conducted of *up.idio* on *SOCSCORE*, *CGVSCORE*, *ENVSCORE* and the seven control variables.

$$up.idio_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \beta_4 Size_{it} + \beta_5 ROA_{it} + \beta_6 SDROA_{it} + \beta_7 Leverage_{it} + \beta_8 MTB_{it} + \beta_9 Liquidity_{it} + \beta_{10} Div.pay_{i,t-1} + \mu_i + \lambda_t + v_{it}$$

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

In total, the model accounted for 45.67% of the variance in *up.idio* and was highly statistically significant.

I fail to reject H_{02} and H_{03}. There is no statistically significant evidence to suggest that the relationship between $CGVSCORE$ and $up.idio$ or $ENVSCORE$ and $up.idio$ is different from zero. Neither *CGVSCORE* nor *ENVSCORE* were statistically significant predictors in the model.

I reject H_{01} in favour of H_{11}. There is statistically significant evidence to suggest that a relationship exists between $SOCSCORE$ and $up.idio$. *SOCSCORE* was a negatively associated highly statistically significant predictor ($\beta = -0.0262$).

These results of the regression results very closely mirror those of *idio* and *ds.idio*. The associations found are very similar to those of *ds.idio*, and both *ds.idio* and *up.idio* closely mirror the associations of *idio* but to a lesser magnitude.

The time-fixed effects were the same statistical significance and direction as those described in the model using the aggregate ESG measure. The strength of the fixed-effect in each year were very similar.

Sub-conclusion, research sub-question 1c

In conclusion, the results of the tests of these hypotheses address research question 1c. There is a significant relationship between *SOCSCORE* and upside idiosyncratic risk but no such relationship between either aggregate ESG, *CGVSCORE*, or *ENVSCORE* and upside idiosyncratic risk.

9.3.4 CSP and idiosyncratic skewness

A multiple linear regression with time-fixed effects was conducted of is on ESG.

$$is_{it} = \beta_0 + \beta_1 ESG_{it} + \lambda_t + v_{it}$$

$$H_0\colon\ \beta_1 = 0$$

$$H_1\colon\ \beta_1 \neq 0$$

In total, the model only accounted for 1.15% of the variance in is but was highly statistically significant.

I fail to reject H_0. β_1 is not statistically significantly different from zero. ESG was not a statistically significant predictor.

The time-fixed effects on the 2009, 2012, and 2015 years were highly statistically significant with a negative association of -0.0827, -0.0961, and -0.0666 respectively. The time-fixed effects in all other years were not statistically significant.

These results indicate that the distribution of residuals was more negatively skewed (i.e., to the left of the distribution) in 2009, 2012, and 2015 cross sectionally across our sample.

A multiple linear regression with random time effects was conducted of is on *SOCSCORE*, *CGVSCORE*, *ENVSCORE*.

$$is_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \omega_{it},\ \ \omega_{it} = \epsilon_t + v_{it}$$

$$H_{01}\colon\ \gamma_1 = 0, H_{11}\colon\ \gamma_1 \neq 0$$

$$H_{02}\colon\ \gamma_1 = 0, H_{12}\colon\ \gamma_1 \neq 0$$

$$H_{03}\colon\ \gamma_1 = 0, H_{13}\colon\ \gamma_1 \neq 0$$

I fail to reject H_{01}, H_{02}, and H_{03}. There is no statistically significant evidence to suggest that the relationship between $SOCSCORE$ and is, $CGVSCORE$ and is, or $ENVSCORE$ and is different from zero. Neither *SOCSCORE* nor *CGVSCORE* nor *ENVSCORE* were statistically significant predictors.

9.3.5 Comparison of results for risk measures

Overall, a comparison of the regression analysis conducted on *idio*, *ds.idio*, and *up.idio* on both the aggregate ESG measure and disaggregate ESG components suggest that these three idiosyncratic risk measures are measuring the same phenomenon in our sample. These measures are not able to distinguish between these two sides of idiosyncratic risk. An analysis of the correlation coefficients shared by the three measures corroborates this evidence.

This finding is corroborated by the lack of statistical significance of the association between ESG measures and *is*. We did not find no evidence that CSP is associated with greater skewness (i.e., to the right of the distribution). If it were the case that CSP decreases downside *idio* and increase upside *idio*, we would expect to see a positive association between the ESG measures and *is*.

Sub-conclusion, research sub-question 1d

In conclusion, the results of the tests of these hypotheses address research question 1d. There is no evidence that the relationships between ESG measures and idiosyncratic risk, downside idiosyncratic risk, and upside idiosyncratic significantly differ.

9.3.6 CSP and abnormal returns

A linear regression with firm-fixed and time-fixed effects was conducted of abnormal returns on ESG.

$$\alpha_{it} = \beta_0 + \beta_1 ESG_{it} + \mu_i + \lambda_t + v_{it}$$

$$H_0\colon \beta_1 = 0$$

$$H_1\colon \beta_1 \neq 0$$

In total, the model accounted for 15.91% of the variance in abnormal returns and was highly statistically significant.

I reject H_0 in favour of H_1. There is statistically significant evidence to suggest that a relationship exists between ESG and α. ESG was a negatively associated highly statistically significant predictor ($\beta = -0.0195$).

These results are consistent with the theory that investors may bear a cost for their social preferences, thus neglecting low CSP firms, as a result of the excess demand for high CSP firms and shortage of demand for low CSP firms (Hong & Kacperzyk 2009; Galema et al. (2008).

The time-fixed effects on the 2007, 2008, and 2011 years were statistically significant with a negative association of -0.0127, -0.01555, and -0.01453 respectively. The fixed-time effect on 2015 was statistically significant with a positive association of 0.0128. The fixed effects on 2004 was marginally statistically significant with a negative association. The fixed effects on all over years were not statistically significant. These results are generally expected and closely expected as the years with negative fixed-effects are those during which the European market saw large negative returns, as discussed in more detailed when discussing the results of the regressions on excess log returns below.

A multiple linear regression with firm-fixed and time-fixed effects was conducted of α on *SOCSCORE*, *CGVSCORE*, *ENVSCORE*.

$$\alpha_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \mu_i + \lambda_t + v_{it}$$

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

In total, the model accounted for 15.91% of the variance in abnormal returns and was highly statistically significant.

I fail to reject H_{01}, H_{02}, and H_{03}. There is no statistically significant evidence to suggest that the relationship between $SOCSCORE$ and α, $CGVSCORE$ and α, or $ENVSCORE$ and α different from zero. Neither *SOCSCORE* nor *CGVSCORE* nor *ENVSCORE* were statistically significant predictors in the model. *CGVSCORE* was however marginally statistically significant predictor ($\beta = -0.0083$).

The time-fixed effects were the same statistical significance and direction as those described in the model using the aggregate ESG measure. The strength of the fixed-effect in each year were very similar.

9.3.7 CSP and returns

A linear regression with firm-fixed and time-fixed effects was conducted of r^e on ESG.

$$r_{it} = \beta_0 + \beta_1 ESG_{it} + \mu_i + \lambda_t + v_{it}$$

$$H_0\colon \beta_1 = 0$$

$$H_1\colon \beta_1 \neq 0$$

In total, the model accounted for 42.56% of the variance in abnormal returns and was highly statistically significant.

I reject H_0 in favour of H_1. There is statistically significant evidence to suggest that a relationship exists between ESG and r^e. ESG was a negatively associated highly statistically significant predictor ($\beta = -0.1148$).

These results indicate that, after controlling for time and firm effects, an increase in 0.01 in yearly ESG predicts a decrease of annual excess log returns of 0.00115 or 0.115%. Equivalently, an increase in one standard deviation of ESG (i.e., 0.239 or 23.9 in the original Asset4 scale) predicts a decrease in *r* of 0.02744 or 2.74%. This is in the context of average annual excess log returns during this time period of 0.015 or 1.50%.

These results closely mirror those of alpha. It appears that the ESG measures have a similar relationship with alpha and *r*. This consistent with an interpretation of alpha as the returns that remain of *r* after controlling for returns attributable to exposure to common risk factors.

As with alpha, these results are consistent with the theory that investors may bear a cost for their social preferences due to excess demand for high CSP firms and shortage of demand for low CSP firms (Hong & Kacperzyk 2009; Galema et al. 2008). ESG may effectively carry be a factor bearing a risk discount.

The time-fixed effects on the 2002, 2008, and 2011 years were statistically significant with a negative association of -0.2371, -0.1648, and -0.0718 respectively. The time-fixed effects on 2007 and 2014 were not statistically significant. The time-fixed effects on all other years were statistically significant with a positive association.

These time-fixed effects results are generally as were expected. An analysis of the yearly log total returns of the EURO STOXX 50 Index, a blue-chip index for the Eurozone used here as a proxy for the broad European market, reveal that 2002, 2008, and 2011 were years with the largest negative returns during the period from 2002 to 2016.

A multiple linear regression with firm-fixed and time-fixed effects was conducted of abnormal on *SOCSCORE*, *CGVSCORE*, *ENVSCORE*.

$$r_{it} = \beta_0 + \beta_1 SOCSCORE_{it} + \beta_2 CGVSCORE_{it} + \beta_3 ENVSCORE_{it} + \mu_i + \lambda_t + v_{it}$$

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

In total, the model accounted for 42.56% of the variance in abnormal returns and was highly statistically significant.

I fail to reject H_{01}, H_{02}, and H_{03}. There is no statistically significant evidence to suggest that the relationship between $CGVSCORE$ and r, $ENVSCORE$ and r, or $ENVSCORE$ and r is different from zero. Neither *SOCSCORE* nor *CGVSCORE* nor *ENVSCORE* were statistically significant predictors in the model. *ENVSCORE* was however marginally statistically significant ($\beta = -0.0552$).

The time-fixed effects were the same statistical significance and direction as those described in the model using the aggregate ESG measure. The strength of the fixed-effects in each year were very similar.

9.4 Robustness checks

9.4.1 Outliers in *r*:

Due to the strength of the relationship between ESG and *r*, additional analysis was conducted to investigate the robustness of the findings.

The total sample contains a small number of very large negative *r* outliers. Despite the large sample size used and choice of firm-fixed and time-fixed effects, it may be the case that these extreme outliers influence the regression results. It may be the case that their extreme negative returns are a function of an error in the *r* calculation in the Datastream database. If they also happen to have high ESG values, this may influence our regression results and complicate inference.

Firstly, an analysis of the ESG performance of the observations with extreme negative *r* outliers was conducted. A subsample of the 150 firm-year observations, from the total sample of 10,827 firm-year observations, exhibit *r* lower than three standard deviations from the mean. This subsample has a mean ESG of 0.5664 and standard deviation of ESG of 0.2153. This is slightly lower than the total sample mean ESG of 0.619 and standard deviation of ESG of 0.239. It appears that the ESG characteristics of the subsample of extreme negative outliers no not different significantly from those of the total sample. This supports the robustness of the main findings.

Secondly, the sensitivity of the regression results to the outliers was tested. *r* was winsorized at the 0.01 and 0.99 level to effectively draw down the top 1% extreme negative and positive outliers to the 99^{th} percentile. The regression was repeated: a linear regression with firm-fixed and time-fixed effects was conducted of *r* on ESG. The same was done at the 0.005/0.995, 0.001/0.999, and 0.001 levels.

The results of the four regressions described above result in a highly statistically significant and negative association. This matches the results of our main regression analysis and provides support for the robustness of the findings on the associations between ESG measures and *r*.

9.4.2 Regression analysis of r_{it}^{e} on ESG with factor betas:

In addition to regressing *r* on ESG and its components, it is also possible to include the estimation of the factor coefficients into the linear regression, in a two-stage regression similar to that of the Fama-MacBeth method and described by Brooks.

In the context of testing the CAPM model, Fama and MacBeth (1973) proposed to handle panel data using a two-step estimation procedure. First, the betas are estimated in separate time series regressions for each firm, and second, for each time period a cross-sectional regression of the excess returns on the betas is conducted. In the second stage, they propose then taking the average of the parameter estimates (i.e., betas) to conduct hypothesis tests. This approach could be adapted in this thesis. However, it is possible to achieve a similar objective using a panel approach (i.e., fixed or random effects) (Brooks, 2014).

ESG could be added to a test of the Fama French five factor model factor betas to examine whether it has still has statistically significant explanatory power after controlling for the explanatory power of the factor betas. A statistically significant positive coefficient would indicate that ESG is a risk factor that earns a risk premium while a statistically significant negative coefficient would indicate that ESG is subject to risk discount.

The Lagrange multiplier test was applied on a pooled linear regression and the results for time effects, firm effects, and two way effects returned χ^2 test statistics of 392,690, 3.1778, 392,690, respectively. The test statistics for individual effects was marginally statistically significant while time and two way effects were highly statistically significant. The pooled linear regression approach is therefore not optimal.

The Hausman test statistic was highly statistically significant. The fixed effects approach is therefore more appropriate. There was a statistically insignificant *DW* test statistic of 2.1058. In contrast, the Breusch-Godfrey/Wooldridge test for higher order autocorrelation returned a highly statistically significant χ^2 test statistic of 30.927. This suggests autocorrelation. The Breusch-Pagan test for heteroscedasticity was conducted. The test returned a highly statistically significant test statistic of 1,507.1, which indicates the presence of heteroscedasticity. None of the explanatory variables exhibited a VIF greater than 10, which supports the assertion that the risk of highly multicollinearity in the explanatory variables is low.

As a result, the following multiple linear regression model with firm-fixed and time-fixed effects was used:

$$r_{it}^{e} = \gamma_0 + \gamma_1 ESG_{i,t} + \gamma_2 \widehat{\beta_{\iota,t,m}} + \gamma_3 \widehat{\beta_{\iota,t,SMB}} + \gamma_4 \widehat{\beta_{\iota,t,HML}} + \gamma_5 \widehat{\beta_{\iota,t,RMV}} + \gamma_6 \widehat{\beta_{\iota,t,CMA}} + \mu_i + \lambda_t + v_{it}$$

$$H_0\colon \gamma_1 = 0$$

$$H_1\colon \gamma_1 \neq 0$$

Where $\widehat{\beta_{\iota,t,m}}, \widehat{\beta_{\iota,t,SMB}}, \widehat{\beta_{\iota,t,HML}}, \widehat{\beta_{\iota,t,RMV}},$ and $\widehat{\beta_{\iota,t,CMA}}$ are the regression coefficients (i.e., factor betas) estimated in the earlier computation of *idio* and alpha for each firm.

The analogous regression and diagnostic tests were conducted for the individual ESG components.

$$r_{it}^{e} = \gamma_0 + \gamma_1 SOCSCORE_{it} + \gamma_2 CGVSCORE_{it} + \gamma_3 ENVSCORE_{it} + \gamma_4 \widehat{\beta_{\iota,t,m}} + \gamma_5 \widehat{\beta_{\iota,t,SMB}} + \gamma_6 \widehat{\beta_{\iota,t,HML}} + \gamma_7 \widehat{\beta_{\iota,t,RMV}} + \gamma_8 \widehat{\beta_{\iota,t,CMA}} + \mu_i + \lambda_t + v_{it}$$

$$H_{01}\colon \gamma_1 = 0, H_{11}\colon \gamma_1 \neq 0$$

$$H_{02}\colon \gamma_1 = 0, H_{12}\colon \gamma_1 \neq 0$$

$$H_{03}\colon \gamma_1 = 0, H_{13}\colon \gamma_1 \neq 0$$

The results using 'HAC' standard errors are contained in an Appendix.

The results above are similar to those of the main regression analyses. ESG remains a negatively associated highly statistically significant predictor ($\beta = -0.0978$) but has decreased in strength by -0.017 or approximately 15%. R^2 rose from 42.56% to 45.11% indicating that the model accounted for more of the variance in abnormal returns. This is to be expected if the factor betas also have explanatory power in the model, which these results indicate is true due to the strength and statistical significance of two of the five factors: market-beta, RMW, and HML.

ENVSCORE remains a negatively associated marginally statistically significant predictor ($\beta = -0.0555$) and has decreased slightly in strength. *SOCSCORE* and *ENVSCORE* both continue to lack statistical significance and have reduced in strength.

These results provide further support for the robustness of the findings regarding the association between CSG measures and *r*.

Sub-conclusion, research sub-question 2a

In conclusion, the results of the tests of these hypotheses address research question 1c. There is significant relationship between ESG and market returns but no such relationship between either *SOCSCORE*, *CGVSCORE*, or *ENVSCORE* and market returns.

9.6 Mediation analysis

Our mediation analysis is simplified by two observations in data: both of the return measures have the same relationship with the ESG measures (i.e., both are only negatively associated with ESG) and each of the different *idio* measures has the same relationship with the ESG measures (i.e., all three are only negatively associated with *SOCSCORE*).

Applying the mediation analysis methodology:

1. It must firstly be established that there is an effect that may be mediated. In our case, this requires regressing returns on CSP. *SOCSCORE*, *CGVSCORE*, and *ENVSCORE* fail this test. There is no evidence of an effect to mediate in the first place. ESG passes this step and moves on.
2. It must secondly be established that the causal variable is correlated with the mediator variable. In our case, this requires regressing *idio* on CSP. ESG does not pass this step. Therefore, there is evidence of a relationship between ESG and returns but no evidence that *idio* mediates this relationship.

Sub-conclusion, research sub-question 2c

In conclusion, the results of the mediation analysis address research question 2c. There is no significant evidence that the relationship between ESG measures and idiosyncratic risk measures mediate the relationship the relationship between ESG measures and market returns.

9.7 Additional discussion

9.7.1 Comparison of ESG measures:

Overall, this analysis supports the assertions that the components of CSP are empirically and conceptually distinct constructs.

This is supported by an analysis of the Asset4 ESG data which reveals that these scores are measures of very distinct types of firm activity.

Analysis at the level of aggregate level obscures the relationships that components share with risk and return. *SOCSCORE* is found to be negatively associated with and predict *idio* risk but no associations were found when analysing ESG, *CGVSCORE*, and *ENVSCORE*, *SOCSCORE*. In contrast, ESG is found to be negative negatively associated with and predict returns but no associations were found when analysing its components *SOCSCORE*, *CGVSCORE*, and *ENVSCORE*. The relationships between these activities and measures are complex and cannot be summarized as simply as "CSP reduces risk and increases return".

9.7.2 Selection bias considerations

There may be a selection bias in our data. It may be the case that only those firms with some minimum level of CSP are even considered for rating in the Asset4 database by Thomson Reuters. In this case, there would be bias in our sample towards more socially responsible firms.

10 Conclusion

Returning to the research question 1: there was no discernable distinction found between the relationship shared by ESG measures and downside idiosyncratic risk in comparison to that between ESG measures and upside idiosyncratic risk.

Returning to the research question 2: no evidence was found that idiosyncratic risk mediates the relationship between ESG measures and market returns.

However, significant results were found during the course of addressing these research questions. There was a significant negative relationship found between *SOCSCORE* and idiosyncratic risk and also *SOCSCORE* and both idiosyncratic risk's downside and upside component. This supports the theory that CSP is associated with lower firm-specific shocks. Secondly, a significant negative relationship was found between aggregate ESG and market returns (both excess returns and abnormal returns). This supports the theory that investors may pay for their social preferences.

Works cited

Aguinis, H. & Glavas, A. (2012). What We Know and Don't Know About Corporate Social Responsibility: A Review and Research Agenda. *Journal of Management, 38*(4), 932-968.

Albuquerque, R, Durnev, A., Koskinen, Y. (2013). *Corporate social responsibility and firm risk: theory and empirical evidence.*(Discussion Paper Series No. 9533) London: Centre for Economic Policy Research.

Alexander, G. J., & Buchholz, R. A. (1978). Corporate social performance and stock market performance. *Academy of Management Journal, 21*, 479-486.

Ang, A., R. Hodrick, Y. Xing, & X. Zhang (2006). The cross-section of volatility and expected returns. *Journal of Finance, 61*, 259-299.

Ang, A., R. Hodrick, Y. Xing, & X. Zhang (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics, 91*, 1-23.

Arrow, K. J. (1974). Limited Knowledge and Economic Analysis. *American Economic Review*, *64*(1), 1-10.

Barnett, M. L. (2007). Stakeholder Influence Capacity and the Variability of Financial Returns to Corporate Social Responsibility. *Academy of Management Review*, *32*(3), 794–816.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management 17*, 771–792.

Bernstein, P. (2005). *Capital Idea: The Improbable Originals of Modern Wall Street*. Hoboken: John Wiley & Sons, Inc.

Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *Review of Financial Studies, 23*(1), 169-202.

Brooks, C. (2014). *Introductory Econometrics for Finance* (3rd ed.). Cambridge: Cambridge University Press.

Sen, S. & Bhattacharya, C.B., (2001). Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility. *Journal of Marketing Research, 38*(2), 225-243.

Bassen, A., Meyer, K., & Schlange, J. (2006). The influence of corporate responsibility on the cost of capital.

An empirical analysis. (Working paper) Hamburg

Bodie, Z., Kane, A., & Marcus, A. (2014). *Investments* (10th ed.). New York: McGraw-Hill Education

Bollen, N. P. B., (2007). Mutual Fund Attributes and Investor Behavior. *The Journal of Financial and Quantitative Analysis*, *42*(3), 683-708.

Bouslah, K., Kryzanowski, L., & M'Zali, B. (2013). The impact of the dimensions of social performance on firm risk. *Journal of Banking and Finance, 37*(4), 1258-1273.

Boutin-Dufresne, F., & Savaria, P. (2004). Corporate social responsibility and financial risk. *The Journal of Investing*, *13*(1), 57-66.

Breusch, T. S. & and Pagan, A. R. (1980). The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *The Review of Economic Studies, 47*(1), 239-253.

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, *52*(1), 57-82.

Carroll, A. B. (1979). A Three Dimensional Model of Corporate Social Performance. *Academy of Management Review*, *4*(4), 497-505.

Chang, K., & Kim I, & Li Y. (2014). The heterogeneous impact of corporate social responsibility activities that target different stakeholders. *Journal of Business Ethics*, *125*(2), 211-234.

Choi, S. (2008). Heteroskedasticity and Autocorrelation [Powerpoint slides]. University College London. Retreived from http://www.homepages.ucl.ac.uk/~uctpsc0/Teaching/GR03/Heter&Autocorr.pdf

Clarkson, M B. E., (1995) A stakeholder framework for analyzing and evaluating corporate social performance. *Academy of Management Review 20*, 92–117.

Coase, R. H. (1937). The nature of the firm. *Economica*, *4*, 386-405.

Crouch, C. (2006). Modelling the Firm in its Market and Organizational Environment: Methodologies for Studying Corporate Social Responsibility. *Organization Studies*, *27*(10), 1533-1551.

Bowen, H. R. (1953). *Social Responsibilities of the Businessman*. New York: Harper & Row.

Dam, L., (2008). Corporate Social Responsibility and Financial Markets. PhD thesis. University of Groningen: Groningen

Davis, K. (1973). The case for and against business assumptions of social responsibilities. *Academy of Management Journal*, *16*. 312–317.

Diemont, D., Moore, K., & Soppe, A. (2016). The Downside of Being Responsible: Corporate Social Responsibility and Tail Risk. *Organization Studies*, *137*(24), 213-229.

Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review 20*, 65–91.

Dorfleitner, G., Halbritter, G., Nguyen, M. (2015). Measuring the level and risk of corporate responsibility – An empirical comparison of different ESG rating approaches. *Journal of Asset Management*, *16*(7), 450-466.

Fama, E. F. & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance, 47*(2), 427-465.

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

Fama, E. F. & French, K. R. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics 116*(1), 1-22.

Fama, E. F. & French, K. R. (2016). Dissecting Anomalies with a Five-Factor Model. *Review of Financial Studies, 29*(1), 69-103.

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

Fombrun, C. and Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal, 33*, 233–258.

Fombrun, C., Gardberg, N.A., & Barnett, M.L. (2000). Opportunity Platforms and Safety Nets: Corporate Citizenship and Reputational Risk. *Business and Society Review, 105*(1), 85-106.

Frederick,W. C. (1995). *Values, nature, and culture in the American corporation*. New York: Oxford University Press.

Freeman, R. E. (1984). *Strategic Management. A Stakeholder Approach*, Boston: Pitman.

Freeman, R. E., & Evan, W. M. (1990) Corporate governance: A stakeholder interpretation. *Journal of Behavioral Economics, 19*(4). 337–359.

Frieder, L., & Jiang, G. J. (2007, July 1). Separating Up from Down: New Evidence on the Idiosyncratic Volatility – Return Relation. AFA 2008 New Orleans Meetings Paper. Retrieved from SSRN:

https://ssrn.com/abstract=970875 or http://dx.doi.org/10.2139/ssrn.970875

Global Sustainable Investment Alliance (2016). *2016 Global Sustainable Investment Review*. Global Sustainable Investment Alliance.

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

Gardberg, N. A., & Fombrun, C. J. (2006). Corporate citizenship: Creating intangible assets across institutional environments. *Academy of Management Review, 31*(2), 329–346.

Godfrey, P. C. (2008). The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Academy of Management Review, 30*(4), 777-798.

Godfrey, P. C., Merrill, C. B., & Hansen, J. M. (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic Management Journal, 30*(4), 425-445.

Gond, J.-P. & Matten, D. (2007). *Rethinking The Business-Society Interface: Beyond the Functionalist Trap*. Nottingham: ICCSR Research Paper Series.

Gond, J.P. & Moon, J. (2011). Corporate Social Responsibility in Retrospect and Prospect: Exploring the Life-Cycle of an Essentially Contested Concept. In Gond, J.P. & Moon, J. (Eds.), *Routledge Major Work on Corporate Social Responsibility*. London: Routledge.

Gujarati, D. & Porter, D. (2014). *Basic Econometrics* (5th ed.). New York: McGraw-Hill Irwin.

Greening, D. W., & Turban, D. B. (1997). Corporate social performance and organizational attractiveness to prospective employees. *Academy of Management Journal, 40*, 658–672.

Greening, D. W., & Turban, D. B. (2000). Corporate social performance as a competitive advantage in attracting a quality workforce. *Business & Society, 39*, 254–280

Heinkel, R., Kraus, A., & Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis, 35*, 431–449.

Harjoto, M., & Jo, H. (2014). Analyst coverage, corporate social responsibility, and firm risk. *Business Ethics: A European Review, 23*(3), 272-292.

Harjoto, M., & Jo, H. (2015). Legal vs. normative CSR: differential impact on analyst dispersion, stock return volatility, cost of capital, and firm value. *Journal of Business Ethics, 128*(1), 1-20.

Hemingway, C. A., & Maclagan, P. W. (2004). Managers' personal values as drivers of corporate social responsibility. *Journal of Business Ethics, 50*(1), 33–44.

Hill, C.W.L. (1990). Cooperation, opportunism, and the invisible hand: Implications for transaction cost theory. *Academy of Management Review, 15*, 500-513.

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

Hong, H., & Liskovich, I. (2017, September 2017). Crime, Punishment and the Value of Corporate Social Responsibility. (Unpublished Manuscript). Retrieved from: https://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?referer=https://www.google.dk/&httpsredir=1&article=1111&context=law_econ

Hull, J. C. (2014) *Options, Future, and Other Derivatives* (9th ed.), Pearson.

Husted, B. (2005). Risk Management, Real Options, Corporate Social Responsibility. *Journal of Business Ethics, 60*(2), 175-183.

Hutton, A.P., Marcus, A.J., & Tehranian, H., (2009). Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics, 94*(1), 67–86.

Jin, L., Myers, S & C., (2006). R^2 around the world: new theory and new tests. *Journal of Financial Economics, 79*(2), 257–292

Jo, H., & Na, H. (2012). Does CSR Reduce Firm Risk? Evidence from Controversial Industry Sectors. *Journal of Business Ethics, 110*(4), 441-456.

Jones, T. M. (1995). Instrumental stakeholder theory: A synthesis of ethics and economics. *Academy of Management Review, 20*, 404–437.

Kempf, A. & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management, 13*(5), 908-922.

Kenny, D. A. (n.d.) Mediation (David A. Kenny). [online] Available at: http://davidakenny.net/cm/mediate.htm [Accessed 15 Jan. 2018].

Kim, Y., Li, H., & Li, S. (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking and Finance, 43*, 1-13.

King, A. (1995). Avoiding ecological surprise: Lessons from long-standing communities. *Academy of*

Management Review, 20, 961-985.

Koch, S. (2010, April 1). Low Risk and High Returns: Evidence from the German Stock Market. (Unpublished manuscript) Retrieved from SSRN: https://ssrn.com/abstract=1632425 or http://dx.doi.org/10.2139/ssrn.1632425

Lintner, J. (1965). The Valuation of Risky Assets and the Selection of Risky Investment in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics, 47*(1), 13-37.

Lee, D.D., & Faff, R.W. (2009). Corporate sustainability performance and idiosyncratic risk: a global perspective. *Financial Review, 44*(2), 213-237.

Lee, M.-T. (2016). Corporate social responsibility and stock price crash risk: Evidence from an Asian emerging market. *Managerial Finance, 42*(10), 963-979.

Luo, X., & Bhattacharya, C.B. (2009). The Debate over Doing Good: Corporate Social Performance, Strategic Marketing Levers, and Firm-Idiosyncratic Risk. *Journal of Marketing, 73*(6), 198-213.

Margolis, J. D. Elfenbein, H. A. & Walsh, J. P. (2009) Does it Pay to Be Good...And Does it Matter? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance (Working paper) (March 1, 2009). Retrieved from SSRN: https://ssrn.com/abstract=1866371 or http://dx.doi.org/10.2139/ssrn.1866371

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance, 7,* 77–91.

Merton, R. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance, 42*, 483–510

McGuire, J. B., Sundgren, A. & Schneeweis, T. (1988). Corporate social responsibility and firm financial performance. *Academy of Management Journal, 31*, 854–872.

Milken Institute [MilkenInstitute]. (2012, May 1). Impact Investing: Managing Risk for Long-Term Returns [Video file]. Retrieved from https://www.youtube.com/watch?time_continue=670&v=Xf5rhLSYpE0

Mitchell, R. K., Agle, B. & Wood, D. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review, 22*, 853–886.

Mitton, T., & K. Vorkink. (2007). Equilibrium Underdiversification and the Preference for Skewness. *Review of Financial Studies, 20*, 1255–88.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, *34*(4), 768-783.

Munk, C. (2016). *Financial Markets and Investments* (this version: September 30, 2016) [PDF]. Retrieved from Copenhagen Business School Investments (LA) CBS Learn site

Noon, M. (2007). The fatal flaws of diversity and the business case for ethnic minorities. *Work, Employment & Society*, *21*(4), 773-784

Oikonomou, I., Brooks, C., & Pavelin, S. (2012). The Impact of Corporate Social Performance on Financial Risk and Utility: A Longitudinal Analysis. *Financial Management, 41*(2), 483-515.

Orlitzky, M., & Benjamin, J. D. (2001). Corporate Social Performance and Firm Risk: A Meta-Analytic Review. *Business & Society, 40*(4), 369-396.

Orlitzky, M., Schmidt, F. L., & Rynes, L. (2003). Corporate Social and Financial Performance: A Meta-Analysis. *Organization Studies, 24*(3), 403-441.

Peloza, J. (2006). Using Corporate Social Responsibility as Insurance for Financial Performance. *California Management Review, 48*(2), 52-72.

Petersen, C. & Plenborg, T. (2012). *Financial Statement Analysis - Valuation - Credit Analysis - Executive Compensation*, Essex: Person Education Limited.

Preston, L.E., & O'Bannon, D.P., (1997). The corporate social–financial performance relationship. *Business and Society 36*, 419-429.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory, 13*(3), 341-360.

Russo, M. V., & Fouts, P. A. (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of Management Journal, 40*, 534–559.

Salama, A., Anderson, K., & Toms, J.S. (2011). Does community and environmental responsibility affect firm risk? Evidence from UK panel data 1994–2006. *Business Ethics: A European Review, 40*(2), 192-204.

Sassen, R., Hinze, A.K., & Hardeck, I. (2016). Impact of ESG factors on firm risk in Europe. *Journal of Business Economics, 86*(8), 867-904.

Scherer, A. G., & Palazzo, G. (2008). Globalization and Corporate Social Responsibility. In A. Crane, D. Matten, A. McWilliams, J. Moon, D. S. Siegel (Ed.), *The Oxford Handbook of Corporate Social Responsibility*. New York, NY: Oxford.

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance, 19*(3), 425-442.

Sharfman, M.P., & Fernando, C.S. (2008). Environmental risk management and the cost of capital. *Journal of Finance, 29*(6), 569-592.

Shrivastava, P. (1995). Ecocentric management for a risk society. *Academy of Management Review, 20*. 118–137.

Spicer, B. H. (1978). Investors, corporate social performance and information disclosure: An empirical study. *Accounting Review, 53*, 94–111.

Thomson Reuters (2013, April). *ASSET4 ESG DATA GLOSSERY* [Excel spreadsheet]. Retrieved from https://uvalibraryfeb.files.wordpress.com/2013/09/asset4_esg_data_glossary_april2013.xlsx

Thomson Reuters (2017, November). *Thomson Reuters ESG Scores*. Thomson Reuters. Retrieved from https://financial.thomsonreuters.com/content/dam/openweb/documents/pdf/financial/esg-scores-methodology.pdf

Treynor, J.L. (1961). Market Value, Time and Risk (Unpublished manuscript).

Ullmann, A. (1985). Data in search of a theory: A critical examination of the relationship among social performance, social disclosure, and economic performance. *Academy of Management Review, 10*, 540–577.

Urdan, T. C. (2010). *Statistics in Plain English* (3rd ed.). New York: Routledge.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*, 171–180.

Waddock, S. A., & Graves, S. B. (1997). The corporate social performance financial performance link. *Strategic Management Journal, 18*, 303–319.

Washer, M. K., & Johnson, R. R., (2013). An Intuitive Examination of Downside Risk. *Journal of Financial Planning,* June 2013, 56-60.

Wang, Q., Dou, J., Jia, S. (2016). A Meta-Analytic Review of Corporate Social Responsibility and Corporate Financial Performance: The Moderating Effect of Contextual Factors. *Business and Society, 55*(8), 1083-1121.

Wang, L., Lin, C., Kao, E., & Fung, H. (2016). Good deeds earn chits? Evidence from philanthropic family controlled firms. *Review of Quantitative Finance and Accounting, 49*(3), 765-783.

Williamson, O. E. (1975). *Markets and hierarchies: Analysis and antitrust implications*. New York: Free Press.

Williamson, O. E. (1985). *The economic institutions of capitalism: Firms, markets, relational contracting*. New York: Free Press.

Zhang, M., & Xie, L., & Xu, H. (2016). Corporate Philanthropy and Stock Price Crash Risk: Evidence from China. *Journal of Business Ethics, 139*(3), 595-617.

Appendix 1 Multicolleanity statistics

CSP and idiosyncratic volatility

	VIF		VIF
SOCSCORE	2.857109	ESG	1.200954
CGVSCORE	1.407703	Size	1.34824
ENVSCORE	2.643504	ROA	1.201044
Size	1.385834	SDROA	1.108318
ROA	1.201051	Leverage	1.037109
SDROA	1.114706	MTB	1.221266
Leverage	1.038571	Liquidity	1.061178
MTB	1.229815	Div.pay.1	1.051
Liquidity	1.083689		
Div.pay.1	1.066086		

CSP and downside idiosyncratic volatility

	VIF		VIF
SOCSCORE	2.857109	ESG	1.200954
CGVSCORE	1.407703	Size	1.34824
ENVSCORE	2.643504	ROA	1.201044
Size	1.385834	SDROA	1.108318
ROA	1.201051	Leverage	1.037109
SDROA	1.114706	MTB	1.221266
Leverage	1.038571	Liquidity	1.061178
MTB	1.229815	Div.pay.1	1.051
Liquidity	1.083689		
Div.pay.1	1.066086		

CSP and upside idiosyncratic volatility

	VIF		VIF
SOCSCORE	2.857109	ESG	1.200954
CGVSCORE	1.407703	Size	1.34824
ENVSCORE	2.643504	ROA	1.201044
Size	1.385834	SDROA	1.108318
ROA	1.201051	Leverage	1.037109
SDROA	1.114706	MTB	1.221266
Leverage	1.038571	Liquidity	1.061178
MTB	1.229815	Div.pay.1	1.051
Liquidity	1.083689		
Div.pay.1	1.066086		

CSP and idiosyncratic skewness

	VIF
SOCSCORE	2.731712
CGVSCORE	1.311264
ENVSCORE	2.567265

CSP and abnormal returns

	VIF
SOCSCORE	2.731712
CGVSCORE	1.311264
ENVSCORE	2.567265

CSP and excess log returns

	VIF
SOCSCORE	2.731712
CGVSCORE	1.311264
ENVSCORE	2.567265

Appendix 2

Regression analysis of r^e on ESG with factor betas

	Dependent variable:	
	(1)	(2)
ESG	-0.0978***	
	(0.0379)	
SOCSCORE		-0.0118
		(0.0349)
CGVSCORE		-0.0293
		(0.0306)
ENVSCORE		-0.0555*
		(0.0309)
market.beta	-0.0607***	-0.0607***
	(0.0071)	(0.0071)
CMA	-0.0056*	-0.0056*
	(0.0032)	(0.0032)
RMW	0.0111***	0.0111***
	(0.0020)	(0.0020)
HML	-0.0141***	-0.0141***
	(0.0031)	(0.0031)
SMB	0.0021	0.0021
	(0.0039)	(0.0039)
R^2 (total)	0.4511	0.4511
Note:		*p<0.1**p<0.05***p<0.01