Faculty of Economics and Business

Thesis MSc Finance

Materiality in non-financial ESG reporting

Name: J.T.T. Dioncre*

Supervisor: Dr. A. Plantinga

Abstract

This paper investigates the relationship between disclosure of material ESG information and corporate financial performance. From a constructed material score based on Thomson Reuters ESG data and materiality taxonomy from the SASB, companies which unexpectedly improve material scores financially outperform companies with lower scores. Using a sample of 2044 firms constituted in 7 worldwide indices from 2003 – 2020, financial outperformance is found across quintile, decile, value and equal- weighted portfolios. Consistent with existing literature in this field, material ESG information proves to be a promising signal for investors looking to improve the signal-to-noise ratio present in traditional ESG investing.

January, 2021

* Faculty of Economics and Business, University of Groningen, Groningen, the Netherlands
E-mail: J.T.T.Dioncre@student.rug.nl Phone: +31 (6) 31544506, Student number S2965615

1. Introduction

Environmental Social and Governmental (ESG) investing is often associated with altruistic motives. However, a global survey by Zadeh & Serafeim, (2018) finds investor embrace sustainability because of the relevance to investment performance and risk management. Addressing all ESG concerns at once is impossible and undesirable. Therefore, the key to success is to focus on ESG risks which are important to a company. This paper explores the concept of materiality and tests whether companies focused on material issues have higher risk adjusted returns.

Materiality is already the norm in financial statements. Information is material if omitting, misstating, or obscuring it could reasonably be expected to influence the decisions that the primary users of financial statements make (IFRS, 2020). 42 percent of investors is dissatisfied with the lack of material sustainability information disclosed (CFA institute, 2017). Between companies active in the same industry, there often is little overlap in ESG reporting, despite these companies face the same regulatory and transitional challenges (Eccles, 2014). Different ESG disclosure is needed between industries. This is explained by an example from the real estate and financial sector.

The real estate sector is responsible for approximately 40 percent of all Green House Gas (GHG) emissions. The biggest challenge faced by this sector in coming years is to align with the Paris agreement. A material topic is to implement mitigation measures that reduce GHG emissions. Additionally, adaptation measures protect assets from the hazardous consequences of global warming. Policies include installing solar panels, mapping of flooding areas, LED-light retrofit programs and installing green roofs. Real estate companies disclose adaptation and mitigation policies to shareholders. This disclosure is material to shareholders because these need to be informed of their progress.

The financial sector faces different sustainability challenges and therefore different material disclosure topics. Robotics and artificial intelligence are opportunities which enhance future productivity and profitability. However, governments will act on automation trends to prevent massive job losses. Another material topic in this industry is customer due diligence. International banks where faced massive anti money laundering penalties because sufficient systems where not in place. Robust customer due diligence systems are important to prevent future scandals. Companies that manage material risks and opportunities are best suited to deal with changes in their operation environment.

Disclosure of sustainability practices is conducted to inform shareholders of the progress made to manage risk and opportunities. Most companies use Global Reporting Initiative (GRI) standards in their ESG reports. The standards offer a very broad scope of ESG topics to chose from. However, not all topics are important to all audiences. GRI recipients include employees, suppliers, society, governments, and shareholders. Employees care about proper health and safety training; suppliers need information about value chain resilience and shareholders disclosure about future ESG risk and opportunities. Most ESG reports are written to disclose information to shareholders. Therefore, companies that want to enhance value from these reports need to focus on information that is important to investors.

In 2018, the Sustainable Accounting Standards Board (SASB) released universal standards for the disclosure of material sustainability information. From a long list of ESG issues, material issues are narrowed down to a short list per industry. The standards were researched and designed to reflect ESG factors that are likely to impact the financial condition and operating performance from companies. In a joined article from 2017, the GRI and SASB chiefs state that both guidelines are not in competition but serve a different purpose for different audiences. The SASB standards serve as an important guide for companies that want to disclose information relevant to shareholders.

Material issues identified by the SASB are chosen because of the value enhancing properties. To check for their success, it is important to study financial outperformance of companies which manage material issues. Kahn et al. (2016) shows companies improving material issues outperform companies that score inferior on these issues. Investment in stocks that improve immaterial issues do not outperform stocks with lower scores. These findings verify material issue topics identified by the SASB can have added value for investors. Only a small amount of ESG issues has an impact on the financial conditions and operating performance of companies.

The purpose of this paper is to examine the relation between material ESG disclosure and financial performance. Material ESG investing has practitioner's relevance as replication studies by institutional investors point out. Additionally, Bloomberg launched a SASB ESG index family in 2019, which includes companies that manage material ESG issues. Using a more global perspective over a longer time horizon than Kahn et al. (2016) , the relationship between material topics and financial performance is further investigated.

Financial performance is evaluated using the Fama and French (2015) five factor model. A material score is constructed using a similar methodology which is used Refinitiv ESG score. Material SASB issues are matched to sustainability datapoints from the Thomson Reuters ESG database. In competitive markets, historic information is incorporated in the share price. Therefore, this study evaluates the change of ESG scores as opposed to score levels. Kahn et al. (2016) use unexpected change as the input for their investment signal. This methodology is used with more caution as unexpected component is not yet grounded in literature.

The main conclusions are similar to other literature in this field. Companies that improve material ESG issues financially outperform companies with inferior ratings on these issues. For 2044 unique firms over a period from 2003- 2020, financial outperformance is found to be robust across decile, quantile, equal and value-weighted portfolios. Therefore, sustainable investors can benefit from differentiating between material and immaterial ESG issues. Investment in immaterial ESG score do not show the same value enhancing effect. No statistical differences are found for strategies which combine both effects. Many companies especially in the early years of the sample do not have ESG data available. This distinction, which is often overlooked, is important to make in ESG research because data is often determined ex-post.

2. Literature review

The main topic of this paper is ESG materiality. However, to understand motivations for investing in ESG issues, general ESG literature is also reviewed. ESG reporting is the main information channel through which shareholders receive sustainability information. Therefore, this literature section discusses how disclosure adds value for shareholders. These reports can be improved towards a point where only material information is disclosed. The focus of this thesis is the disclosure of this material information.

2.1 General ESG investing and ESG reporting

Two views on the link between Sustainability and corporate financial performance (CFP) exist. In one view, ESG investments are more expensive and therefore must destroy shareholder value. Managers are not able to balance the scoreboard for all stakeholders because they end up being shorthanded. In a competitive environment diversified focus could have serious consequences (Jensen, 2010). In the second view, ESG investments increases value because stakeholder risks are handled and mitigated. This protects a company against transitional liabilities and future unknown risks. Adams (2019) finds the main reasons for companies to

incorporate ESG practices is to enhances the corporate strategy, decreases reputational damage, and satisfies investor demand.

A large meta study by Friede et al. (2015) finds 90 percent of over 2000 studies identify a non-negative ESG-CFP relationship. Generating positive alphas with ESG portfolios is difficult because returns depend on the overlapping effects of systematic and idiosyncratic risks (Campbell et al. 2001; Luo & Bhattacharya 2009). Investing in ESG score levels is associated with long-term value driven motivations present in traditional ESG beliefs. ESG score change is yields short-term profits, whereas these effects diminish in the long run (Derwall et al., 2011). Additionally, Brammer and Wellington (2008) find companies having low ESG scores have higher risk adjusted returns in the short-run and high ESG performers are associated with superior long-run performance.

Because investors posses' different horizons, three predominant ESG investment strategies emerge (Nagy, et al 2016). An ESG-exclusion strategy excludes worst performers per industry but diversifies its investment across industries. Excluded companies have the largest probability to damage long-term returns but the portfolio is still diversified. An ESG-tilt strategy shifts the investment towards companies with higher ESG scores. Both strategies assume sustainability scores are linked to long-term returns and future losses can be avoided (value driven). An ESG-momentum strategy evaluates the change in company ESG scores. This strategy also relies on the avoidance of future liabilities but assumes these are quickly priced-in by the market. Nagy et al. (2016) compares both strategies and finds both ESG-tilt and ESG-momentum outperform the world index. Investment using an ESG-momentum strategy shifts the portfolio towards stocks with positive price momentum. The ESG-tilt strategy pushes the portfolio towards less volatile stocks.

Companies have an increasing impact on society and their operating environment. The first sustainability reports were released to mitigate these concerns and came as answer to pressure from environmental organizations. The most used forms of ESG disclosure are integration in the yearly report or a stand-alone report. Positive market valuations associated with ESG disclosure are not influenced by the type of reporting used (Mervelskemper & Streit, 2017). The Global Reporting Initiative (GRI) is the best-known organization that develops a set of guidelines for producing sustainability reports worldwide. Around 75 percent of the world biggest 250 companies use the reporting framework. The standards offer important disclosure topics to both internal and external stakeholders. These stakeholders include society, suppliers, governments, creditors, employees, managers, shareholders, and customers.

Self-reported, unaudited, and unstandardized sustainability reports tend to showcase companies in the best light. A solution to limit ESG information disclosure is to force public companies to release ESG information (Lydenberg el al.,2012). In the UK and Germany, companies listed on major stock exchanges must report non-financial issues relevant to their business. To obtain government funding in the US, companies must disclose certain non-financial ESG items in their 10-K filings. From March 2021, partial ESG disclosure becomes mandatory in the European union. Large companies participating in financial markets with over 500 employees must disclose parts of their ESG information (EU Disclosure Regulation 2088, 2019)

2.2 Origin of Material ESG

GRI standards do not meet the very specific but diversified needs of its potential audiences. Since ESG reporting is primarily written for shareholders, the focus of should be on what is important to shareholders. Brown (2011) states the usage of GRI standards by its intended audience has been rather low. Shareholder activism organizations, institutional investors and socially responsible investors treat company ESG reports as supplemental material rather than their main research and strategy source. Despite the efforts for standardization, ESG disclosure with GRI standards has limited added value for cross company performance comparing (Brown et al., 2009).

Companies active in the same industry face similar regulation and comparable competitive advantages. Their reporting is therefore expected to converge to a single format. Eccles (2012) finds much deviation between topics and scope that companies report on. For example, in the airline industry fuel prices are a material issue. The disclosure topics range from R&D biofuel investments, fuel conservation practices, climate change programs and carbon regulations. The scope on these same topic's ranges from no disclosure to robust quantitative metrics. Standardization of material topics is needed before investors can make a peer-to-peer comparison between investment decisions.

Lydenberg et al. (2012) develop a six-step procedure to determine key sustainability performance indicators (KPI) within an industry. Material topics are selected from a broad range of issues based on, regulatory drivers, peer-based norms, stakeholder concerns and opportunities for innovation. The most important topics in the airline industry are environmental issues, but in the banking industry these are governmental. Disclosure should differ between industries to reflect important topics to that sector. Value chain membership and

geography location are other criteria to determine material topics, but these have received less attention.

2.3 The Sustainability Accounting Standards Board (SASB)

Material ESG disclosure has led to the establishment of the SASB which was founded in 2011. Its mission is to establish and improve industry specific disclosure standards for ESG topics that are important to shareholders. The work by Eccles, (2012) and Lydenberg et al. (2012) have been important attributions to identify material issue topics. Universal applicable indicators are narrowed down to items which are material. Sustainability reports should prioritize material information and switch vague language to quantitative verifiable metrics that are useful for investors.

To determine value adding sustainability issues, a relevant peer group is determined first. This industry classification system, groups companies by their sustainability threats and opportunities. The process outlined by Lydenberg et al. (2012) determines important KPI's relevant to the industries. Issues are identified, assessed, developed, proposed revised and released on an issue-by-issue basis. Furthermore, topics are always evidence based and/or market informed (SASB conceptual framework, 2020). Feedback on the implementation and functioning of the standards is provided by over 150 market informed industry experts.

According to the SASB, disclosure on material topics characterizes the company's positioning with respect to business-critical sustainability issues. Whether issues create long-term value is assessed by three financial metrics; (1) Revenue and cost effects, that could be gained by operational efficiencies that arise from sustainable investments. An example is the decrease in electricity or water costs. (2) Asset and Liability effects, on assets which a company possesses. For example, agricultural land which can be impaired by water scarcity. (3) Market valuation effects, where the improvement of a company's investor assigned risk profile is changed by disclosure. The difference between traditional ESG and material ESG is the number of issues that are important. The SASB's allocation process is designed to narrow down issues to be value enhancing in a particular industry.

2.4 Financial outperformance

Financial outperformance was first investigated by Kahn et al. (2016) and their findings are promising. Firms with strong ratings on material sustainability issues outperform firms with inferior ratings. Investment in stocks that are superior on immaterial issues do not outperform stocks with worse ratings. Significant financial outperformance is found using a calendar-time

portfolio stock return and firm-level panel regressions. The study uses data the MSCI KLD 400 Social Index, which has only coverage of the US. Binary strength and concern flags from a wide range of sustainability issues are mapped to issues from the SASB materiality map.

Replications studies based on Kahn et al. (2016) discussed below are underwritten by larger investors. Therefore, these papers could be interpreted as less academic in nature. Steinbarth (2018) underwritten by Russel Investments, builds an improved material ESG score. This score is superior at predicting future financial performance. The correlation between known return drivers and their material score are larger. For sustainability investors, it is important to distinguish between material and immaterial issues. Much of the signals feeding in traditional ESG scores are not important for value creation purposes.

Kotsantonis (2019) underwritten by The Global Alliance for Banking on Values (GABV), tests financial performance in the commercial banking industry. Sustainable banks that consistently score high on material ESG issues delivered higher risk-adjusted returns. Banks that score inferior on material issues have lower returns. This relationship was not found for immaterial ESG issues. Outperformance is only found after 2013, in the period 2007-2013 portfolios based on high and low material scores behave similarly. Possible explanations are ESG issues have become more important to investors, or improved ESG data quality and coverage enhances measurability of material items. Industries specific research like this, brings additional insights in the workings of material ESG.

A working paper by Heijningen, (2019) underwritten by RobecoSAM derives material ESG issues from internal research. In line with the previous materiality studies, material scores are better predictors for financial performance than immaterial scores. Like traditional ESG and findings from Derwall et al. (2011), to predict long-term financial performance material score levels are better. Other research uses score change to evaluate short-term performance implications. The methodology used to build a material score is different. Weights are placed on three letters of ESG based on their importance to that industry. With this methodology, many datapoints are used in the score calculation. Since only a small list material topics are important per industry, this moves away from true materiality.

2.5 Broader materiality

Financial outperformance originates from a long list of difficult to measure risk management practices. Therefore, material ESG research should measure material effects broader. Henisz and McGlinch (2019) link material Truvalulabs ESG data to an increase in credit risk.

Companies with worse material ESG criteria defined by the SASB, experience more media reported negative credit events. These events take place less for firms with stronger material ESG performance. Therefore, improved performance on material ESG issues offers investors fewer surprises and less volatility. Managers that posses a long-term orientation can benefit from investment in material issue topics.

Henriksson et al. (2019) finds investors value S&P 500 companies with good, bad, and neutral material ESG scores differently. The median Market-over-Book ratio for firms with good material ESG practices is 3.48. For firms with bad scores this ratio is 2.59. This discrepancy is persistent throughout the sample from 2004 – 2018. Therefore, if a company wants to improve its market valuation using ESG practices, it should invest in material ESG issues first. Companies score unfavorable on all SASB material issue have decreased from 218 in 2004 to just 22 in 2018. Firms have recognized the importance of these material ESG issues to investors.

Material ESG is also gaining traction within the investment community. State street global advisor put material investing in practice with the R-Factor™. This score is based directly on the SASB material framework. Data from 4 global data providers and 91 metrics is combined into a single score per firm. The R-Factor™ helps companies build a roadmap for items which need to be disclosed and implemented. For shareholders, the score helps to identify the main engagement topics.

Bloomberg used the R-Factor™ to launch a SASB ESG index family in September 2019. A ESG tilt strategy is used to identify companies which have the largest exposure to the R-Factor™. Separate indices for growth, value, dividend yield stocks as well as corporate bonds exist. According to the SASB, the Bloomberg SASB ESG index family is an innovative example of bringing its vision of materiality based ESG investing to life. Furthermore, the indices enhance capital flows to companies with superior material ESG practices.

2.6 Research outlook

Disclosure as opposed to performance improvements is more important in the Thomson Reuters dataset. Therefore, the hypothesis is adapted to encompass the specific characteristics of the dataset and focus mainly on the disclosure of material ESG data. Financial performance is evaluated using abnormal stock returns unable to be explained by known risk factors. The difference between portfolios based on high and low material disclosure is tested against the alternative that no difference exists.

Hypothesis 1: Unexpected improvements of material ESG disclosure have a positive effect on financial performance

Material ESG issues are selected with a business centric viewpoint in mind. The SASB claims, material issues enhance long-term shareholder value which results in positive risk adjusted returns. Financial performance reflects the improvement companies make in material ESG issues. Because markets are competitive, future value effects are incorporated in the share price immediately. Therefore, material ESG improvements have an immediate effect on financial performance. Kahn et al. (2016) finds evidence of this hypothesis in the US with a different ESG dataset. Using a different dataset, claims from the SASB are verified more extensively.

Hypothesis 2: Unexpected improvements of immaterial ESG disclosure have an insignificant effect on financial performance

Companies reporting immaterial ESG metrics, disclose many issues which are not value enhancing in the SASB material taxonomy. Immaterial ESG is comparable to traditional ESG, because both disclose many topics which are not material. The ESG-CFP relationship has been widely investigated and most studies find a non-negative link (Friede et al. 2015). Therefore, immaterial ESG is disclosure is expected to have the same non-value enhancing effect. The implication is sustainability reporting do not need to focus on immaterial issues. Additionally, this hypothesis strengthens confidence in the first hypothesis.

Hypothesis 3: Firms improving material ESG scores financially outperform firms improving other ESG scores

Sustainability reports contain both material and immaterial information. Companies improving on all issues are score high on both material, immaterial, and Refinitiv scores. Direct comparison between improvements on several scores is investigated in this hypothesis. The Refinitiv score already incorporates a methodology to identify material ESG issues. Therefore, the direct comparison between both scores interesting.

Hypothesis 4: Firms improving material and decreasing immaterial scores outperform firms decreasing material and improving immaterial scores

Given the first and second hypothesis, from an optimization perspective implementing both strategies is further value enhancing. Double sorting portfolios captures the combined effect of both hypothesis. Invest in material and not in immaterial is most productive because less

recourses wasted non-value enhancing activities. Additionally, immaterial ESG disclosure is problematic because investors read over material information.

3. Methodology

This paper examines the relationship between changes in sustainability scores and changes in stock price. The methodology is outlined in three chapters of this section and is similar to Kahn et al. (2016). First, materiality scores are calculated like Refinitiv ESG scores, but with material taxonomy from the SASB. Second, unexpected score change is determined with a cross-sectional regression model based on publicly available firm fundamentals. Third, portfolios from unexpected change in material ESG scores are tested for abnormal financial returns.

3.1 Material ESG score

How well a company manages business-centric sustainability issues is determined by a material score. This material score is constructed from the individual datapoints that make up the Refinitiv ESG score. Therefore, the material score resembles this score but different weights and datapoints are used in the construction. This construction process is summarized in three steps:

1. Identify material ESG metrics and their weights using the SASB sector materiality map

    Appendix 2 shows the complete SASB materiality map

2. Match material ESG metrics to their relevant ranked Thomson Reuters datapoint

    Appendix 3 shows the datapoint chosen for each sector

3. Aggregate datapoints scores for all material issues to a single score per company per year

Disclosure datapoints are binary by design. Either a company has or does not have implemented a policy on a certain issue. Therefore, these variables are converted to a numeric ranked position. The eventual score is influenced by the number of other companies which have implemented a policy. This feature is important to determine how well a company is doing compared to its peers. The Refinitiv use this exact methodology which is shown in the following equation:

$$r_{f,u,t} = \frac{\#\ worse\ disclosure_{u,t} + \frac{\#\ same\ disclosure_{u,t}}{2}}{\#\ within\ sector_t} \quad (1)$$

Where $r_{f,u,t}$ is the ranked position of firm f, for issue u at time t, # worse disclosure is the number of companies without disclosure or implementation of a policy, # same value is the number of companies which have the same disclosure level, and # within sector refers to all companies in a sector at time t. A score of 100 percent is awarded when there is one company with disclosure and many companies without.

Material issues are classified following the taxonomy from the SASB. Using the same procedure as Kahn et al. (2016) and Steinbarth (2018), issue categories are mapped to individual datapoints. Both papers map similar datapoints to industries active in the same sector. This paper therefore maps datapoints to the eleven SASB sectors instead of the industries. Additionally, mapping to sectors reduces noise and results in more consistent usage of datapoints. The number of industries with material issues determines the materiality level for a sector. Sector weighted materiality is shown in the following equation:

$$m_{u,s} = \frac{i_{u,s}}{n_s} \tag{2}$$

Where $m_{u,s}$ is the weighted materiality of issue u in sector s, $i_{u,s}$ is the total number of industries that have issue u material in sector s, and n_s refers to the total number of industries in sector s. The eventual material score is the weighted materiality multiplied by the matched disclosure rank for all material issues. This yields a score per firm for all years in the sample. The immaterial score is determined using the same methodology but with non-material issues. This is shown in equation 4.

$$score_{i,t} = \sum_{u=1}^{n} r_{f,u,t} * m_{u,s} \tag{4}$$

Where $score_{i,t}$ is calculated for each firm i in year t. The ranked position is multiplied by the weighted materiality level for issue 1 to n. Consistent with the Refinitiv ESG score, the sum of the potential score is always 100 percent. With 10 equal weighted issues in a sector, each item is attributes 0.10 to the overall material score. The result is a modified Refinitiv ESG score which only contains material datapoints and material weights per sector. Characteristics from the constructed material and the traditional score are directly compared in appendix 4.

3.2 Unexpected scores

Sustainability investment levels are already incorporated in share price. Therefore, the change in score used for the construction of portfolios. Change is the difference in score compared

to t_{-1}. To mitigate concerns about correlated firm characteristics with the investment signal, like Kahn et al. (2016), changes in material scores are orthogonalized to isolate their unexpected component. Unexpected sustainability change is found using a cross-sectional regression model with fundamental firm characteristics as explanatory variables. This regression is run yearly so unexpected scores are not based on historic data. Unexpected score is the residual from the following equation:

$$\Delta\, score_{i,t} = b_1 + b_2 \Delta Size_{i,t} + b_3 \Delta MTB_{i,t} + b_4 \Delta ROA_{i,t} + b_5 \Delta Leverage_{i,t} + b_6 \Delta Capex_{i,t} + f_s + e_{i,t} \quad (5)$$

Where $score_{i,t}$ is the materiality score for firm i at time t from equation 4. Change in firm characteristics is the value compared to t_{-1}. Sector fixed effects (f_s) are used instead of industry fixed effects as the differences in performance are expected to be higher between sectors compared to industries. Kahn et al. (2016) solve lumpy distributions in material scores using a multivariate regression model. Material scores in this paper consider the relative disclosure compared to other firms. Therefore, more variation exists between these scores. Portfolios already have approximately the same size when results are mean adjusted by their sector. Still, a multivariate regression model is tested, but this did not yield different results.

The residuals from equation 5 are used as investment signals to construct top and bottom portfolios. Firms that rank in the upper (lower) decile are placed in the top (bottom) portfolio. For quintile portfolios the same procedure is used. Unexpected change in material ESG scores is explained by a firms change in size, value, return on asset, leverage, and capital expenditures. These variables represent fundamental characteristics of a firm in terms of size, growth opportunities, valuation, financial structure, and investment profile.

Double sorting portfolios is not possible with Thomson Reuters data, as not enough firm score both in the top material and in the bottom immaterial portfolios. Van Heijningen (2019) only found 20 cases per year where firms have improved on materiality and decreased on immaterial issues. Alternatively, the high materiality low immateriality (HMLI) portfolio is found by subtracting the normalized immateriality score from the normalized materiality score. A LMHI portfolio is found by subtracting the normalized materiality score from the normalized immateriality score.

3.3 Performance evaluation

New ESG information is not incorporated in share price immediately, and not all companies release their sustainability reports at the beginning of the year. Therefore, an event window,

consistent with Kahn et al. (2016), is chosen that resembles a realistic trading strategy. Publicly available information from January is used to construct investment portfolios from April to March next year. Refinitiv first released ESG information in January of 2003. Because the change material score is used, the first portfolio period ranges from April 2004 to March 2005.

Abnormal returns are found with the Fama and French (2015) monthly cross-sectional regression model. This model is not used frequently in ESG literature, because of the interactions with other (older) factors and lack of a human capital factor. However, Blitz and Fabozzi (2017) find the five-factor model can compensate for the inclusion of sin stocks (tabaco, gambling, etc.). The reputation risk premium associated with the investment in sin stocks, disappears with the addition of two factors. The five-factor model is shown in the following equation:

$$r_{i,t} - r_{f,t} = a_i + \beta_i\left(r_{m,t} - r_{f,t}\right) + s_i SMB_{i,t} + s_i HML_{i,t} + r_i RMW_{i,t} + c_i CMA_{i,t} + u_{i,t} \quad (8)$$

Where the dependent variable $r_{i,t} - r_{f,t}$ is the excess return of firm i over the risk-free interest rate. This excess return is explained by the return of the market portfolio over the risk-free rate $r_{m,t} - r_{f,t}$. As well as by $SMB_{i,t}$ the return on a diversified portfolio of small stocks minus the return a diversified portfolio of big stocks, $HML_{i,t}$ the difference between the returns on diversified portfolios of high and low B/M stocks, $RMW_{i,t}$ the difference between the returns on diversified portfolios of stocks with robust and weak profitability, and $CMA_{i,t}$ the difference between the returns on diversified portfolios of the stocks of low and high investment firms. The residual $u_{i,t}$, is tested for heteroskedasticity, autocorrelation and stationary conditional variance. The adjusted R-squared, which adjusts for the number of the terms in the model, is the explanatory power of the model.

Abnormal returns are left unexplained by known systemic risk factors. Therefore, these returns can be interpreted as alpha attributed to material ESG trading strategy. Portfolio performance is evaluated by the abnormal returns from the top portfolio minus the bottom portfolio. Proceeds from going short in the bottom portfolios can be used to purchase shares from the top portfolio. Because this strategy does not require an initial investment, it can be perceived as risk-free.

3.4 Robustness tests

Computing decile, quintile, equal and value-weighted portfolios is the first approach to add validity to the findings. Raw returns and alphas from the Fama and French (1993) three-factor model and the four-factor (Carhart 1997) model are tested for significant differences. The

momentum factor is important as an ESG change investment strategy tilts a portfolio to firms with higher price momentum (Nagy, 2016). Additionally, an alternative unweighted materiality matrix without equation 2 is tested. As opposed to, the weighted level of industries that have the issue material. In this matrix, all issues receive the same weight in a sector.

The hypothesis is separately tested by the exclusion of pollutive and human capital-intensive industries. Polluting firms are more sensitive to environmental legislation, criticism, reputational risk, and environmental disasters (Derwall et al., 2005). High performers on material ESG issues are less effected by these negative risks. Human capital-intensive industries in general are associated with higher ESG scores. Companies in these sectors invest more in human resources, human capital management, recruitment, and retainment (Lo & Sheu, 2007). Company size has the largest correlation with company ESG score. Therefore, the exclusion of the largest half of the companies from the sample is tested separately. Kotsantonis (2019) finds material ESG investment outperformance only in the second half of their study. Therefore, the main analysis is repeated using a split time window.

4. Data and descriptive statistics

4.1 Sample and ESG data

Sustainably datapoints are downloaded from the Thomson Reuters ESG database. Datapoints from this database are compiled in 10 pillar which together form the Refinitiv ESG score. Sustainability information is scraped from annual reports, stock exchange filings and news sources. All datapoints are processed by content research analysis, checked by postproduction quality checks, independent audits, and management reviews. New information is updated on a two-week basis. Compared to other datasets, which are discussed in appendix 1, Thomson Reuters has less focus on performance and more on disclosure. Because the primary objective from the SASB is material ESG disclosure, this dataset is a good fit. Furthermore, between the dataset used by Kahn et al. (2016) and the Thomson Reuters dataset there is only a 0.4 correlation (Kerber & Flaherty, 2017).

Thomson Reuters gives material topics more importance in the Refinitiv ESG score. This is done using a dynamic magnitude matrix where important topics are given more weight. Datapoint weights are determined by the relative disclosure of industry peers. This methodology has little resemblance to materiality from the SASB, which base material issues on academic literature. Only under the assumption that companies disclose just material data do both methodologies have the same outcome.

Table 1. Companies in sample by their index

Index	Country	Current constituents	Past constituents	Total
S&P 500	USA	517	191	708
FTSE 250	UK	254	263	517
NASDAQ 100	USA	104	201	305
FTSE 100	UK	103	242	345
CAC 40	France	40	33	73
DAX	Germany	30	27	57
SMI	Switzerland	20	19	39
Total		1068	976	2044

Table 1 shows the total number of firms included in the sample by their respective index. Past constituents refer to companies that are not currently listed in the index but have been in the past. Total is the sum of the current and past constituents

Only companies for which Thomson Reuters states to have data ranging back to 2003 are included in the sample. This includes all constituents from the S&P 500, FTSE 250, NASDAQ 100, FTSE 100, CAC 40, DAX, SMI from 2003 to 2020. Other indices are added later, but these do not have data available for a sufficient period needed to test financial outperformance. From these constituents, 19 firms have been removed due to missing return data. Table 1 shows a complete overview of the sample which contains 2044 unique firms. Kahn et al. (2016) use a sample of only US firms dating back to 1992. A more global sample is used in this study of which only half of companies is US based. This sample is good proxy for the investment universe of the average investor.

Not all companies have ESG data available for the full period. Table 2 shows in 2003 only 364 firms have data available compared to 1023 which do not. Henriksson (2019) finds similar amounts of missing data while using the Thomson Reuters database. Because ESG scores are determined ex-post, not all information is available for the full period. Other literature omits firms missing ESG datapoints from their sample. This has the potential to create a survivorship bias towards companies that still exist. A separate analysis for companies' whit data missing data is therefore included in this paper.

Table 2 shows from 2044 companies in the sample, 138 companies could not be linked to a SASB sector. The eleven sectors from the SASB materiality map are not a complete list of all industries. Of the unindexed companies, 135 are collective investments and 3 are holding companies. These companies are excluded from the sample because it is not possible to determine a material score. A complete overview of all sectors included in the SASB materiality map can be found in appendix 2.

Table 2. Listings in sample by year

Year	In sample	No SASB Industry	No ESG data	Total with ESG data
2003	1481	94	1023	364
2004	1496	94	866	536
2005	1504	99	657	748
2006	1497	102	472	923
2007	1485	108	428	949
2008	1452	109	388	955
2009	1419	108	332	979
2010	1419	108	300	1011
2011	1407	107	264	1036
2012	1390	107	231	1052
2013	1392	107	235	1050
2014	1410	109	262	1039
2015	1410	111	262	1037
2016	1380	94	166	1120
2017	1362	94	117	1151
2018	1346	99	99	1148
2019	1319	102	76	1141
Individual	2044	138	1023	1781

Table 2 shows the total amount of firms which are included in the sample and have ESG data available. Total with ESG data is the number of firms in sample minus firms with a missing industry, minus firms with missing ESG data.

4.2 Calculation of materiality score

The material score is constructed from individual datapoints that make up the Refinitiv ESG score. Datapoints are transformed to ranked positions to account for the relative industry disclosure. For 10 of the 36 datapoints Thomson Reuters has already calculated a ranked position, for 26 this is done using equation 1. Only datapoints which have at least a 10% disclosure and implementation rate have been used. The Refinitiv ESG score is constructed from 178 individual datapoints which do not vary across sectors. The materiality score uses 26 issue categories of which 5 are on average material to a sector.

Consistent with Kahn et al. (2016), individual SASB issue categories are matched to datapoints from the ESG database. Although there is no complete overlap, Thomson Reuters offers 450 ESG datapoints to chose from. SASB issue categories mean different topics in different industries. As an example, the issue category product design & lifecycle management in real estate industry means: “Discussion of approach to measuring, incentivizing, and improving sustainability impacts of tenants”. However, for the Consumer goods Industry this issue means:

"Description of strategies to reduce environmental impact of packaging throughout its lifecycle". Therefore, no single datapoint per issue can be chosen across sectors. In total 36 datapoints are used to quantify material topics. A complete overview of SASB sectors matched with Thomson Reuters ESG datapoints can be found in appendix 3.

Table 3 presents descriptive statistics of material and immaterial scores as well as the Refinitiv ESG score. The mean material score is 28.50, which is lower than mean Refinitiv ESG score which is 45.75. This difference originates from the implementation rate of datapoints used in both scores. The skewness of the materiality and immateriality score is higher because less datapoints are used. Positive mean deltas of 2.37, 2.36 and 2.10 show that companies have improvement sustainability practices over the years. Companies have not improved their material score more than their immaterial score. Appendix 4 shows histograms on the material and Refinitiv ESG scores.

Table 3. Descriptive statistics ESG scores

	Mean	Std. Dev.	Min	Max	Skew.	Obs.
Material	28.50	19.75	0.00	81.42	0.33	16 537
Δ Material	2.37	8.53	-73.95	77.35	1.21	14 978
Immaterial	24.83	18.42	0.00	85.49	0.50	16 492
Δ Immaterial	2.36	7.87	-64.62	62.19	0.86	14 193
Refinitiv ESG	45.75	20.23	0.63	95.07	0.17	17 115
Δ Refinitiv ESG	2.10	8.23	-81.96	65.24	0.49	15 927

Table 3 presets the summary statistics for the score values. Δ variables represent the absolute change compared to 12 months ago. TR ESG is the traditional Thomson Reuters ESG score. Material and Immaterial are calculated using equation 3

4.3 Unexpected scores and portfolio construction

Table 4 presents summary statistics for the explanatory variables used to determine unexpected improvements. Company characteristics are collected, simultaneously with the ESG data points, yearly in January. The downsides of a broader sample are less consistent availability of datapoints. For P/B, ROA, Leverage and Capex there are less observations due to missing data. By design of the methodology, missing fundamental values are filled with the mean value. Companies lacking all explanatory variables therefore receive their actual change in score as investment input. The average market capitalization is 17.4 billion. Compared to all companies, companies in the sample are relatively large.

Table 4. Descriptive statistics of the explanatory variables

	Mean	St. dev	Min	Max	Obs.
Size	9.76	0.70	6.09	12.01	21 485
ΔSize	0.00	0.94	-8.07	8.38	20 127
P/B	1.78	0.51	-3.96	3.98	20 956
ΔP/B	0.00	0.50	-6.71	7.04	19 654
ROA	0.07	0.05	-0.30	0.30	15 986
ΔROA	0.00	0.05	-0.50	0.49	13 917
Leverage	0.13	0.11	0 .00	2.03	21 071
ΔLeverage	0.00	0.08	-1.32	1.32	19 290
Capex	0.12	0.10	0.00	0.99	10 575
ΔCapex	0.00	0.11	-0.97	0.97	8 354

Table 4 presets the summary statistics for the explanatory variables used for the construction of unexpected signals. Δ variables represent the absolute change compared to 12 months ago. Size = The natural logarithm of market capitalization / 1 million; P/B = Market value over book value of equity; ROA = Total income over the average of total assets of the current and previous year; Leverage =Total debt over the average of total assets of the current and previous year; Capex = Total capital expenditures over the average of total assets of the current and previous year

Table 5. Coefficients unexpected changes in scores

	ΔMaterial Score		ΔImmaterial Score		ΔRefinitiv	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Intercept	1.2651	8.62	1.5256	9.79	1.1893	8.05
ΔSize	-0.7091	-9.03	-1.0460	-12.60	-1.0747	-13.64
ΔP/B	0.0178	0.43	0.0837	2.00	0.0596	1.50
ΔROA	-3.4498	-4.21	-2.3770	-2.74	-3.3608	-4.08
ΔLeverage	4.2144	6.23	0.3641	0.51	1.7246	2.54
ΔCapex	-0.2217	-0.51	-2.2250	-4.74	-1.9980	-4.49
Sector Fixed effects	Yes		Yes		Yes	
Number of years	17		17		17	
Adjusted R-squared	2%		1%		1%	

Table 5 presents the time-series average of estimated coefficients from yearly cross-sectional regressions. The first column uses the change in the total Material score as the dependent variable. The second column uses the change in the Immateriality Index as the dependent variable. The third column uses the change in the Thomson Reuters ESG score as the dependent variable

Table 5 presents the coefficients from the explanatory variables used to determine unexpected change in scores. The average of 17 regressions is presented as the cross-sectional model is run yearly. Sector fixed effects are used as Kahn et al. (2016) finds industry fixed effects does raise the explanatory power of the model. The adjusted R-squared of 2 percent means the model has little explanatory power. Low values are common for models that predict change in values. The adjusted R-squared for the same model that predicts score level is 25 percent. A significant t-

statistic of 9.03 is found for the change in size coefficient. This relationship with size is expected as larger firms have more recourses available to accurately report their data (Drempetic et al., 2020).

Table 6. Correlation matrix of score variables, unexpected levels, and explanatory variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Material (1)	1.000									
Unexpected Material (2)	0.862	1.000								
Immaterial (3)	0.649	0.536	1.000							
Unexpected Immaterial (4)	0.544	0.631	0.849	1.000						
Refinitiv (5)	0.770	0.635	0.735	0.616	1.000					
Unexpected Refinitiv (6)	0.624	0.729	0.594	0.707	0.864	1.000				
Size (7)	0.375	0.005	0.422	0.008	0.452	-0.002	1.000			
P/B (8)	0.005	0.010	-0.023	0.007	-0.006	-0.009	-0.253	1.000		
ROA (9)	-0.039	0.003	0.003	0.004	-0.039	0.003	0.171	-0.162	1.000	
Leverage (10)	0.005	-0.004	0.043	-0.004	0.003	-0.013	0.022	-0.016	-0.242	1.000

Tables 6 shows the correlation between unexpected material score and unexpected Refinitiv score is 0.73. This is higher than the 0.63 correlation that Kahn et al. (2016) finds. A negative or very low correlation would be surprising since these scores measure different areas of sustainability. A very high correlation would suggest there is little added value to rebuilding a material score from the ground up.

Between the material and immaterial score, the correlation is 0.63 which is higher than the 0.30 correlation that Kahn et al. (2016) finds. Positive correlation is expected as the datapoints used are the same but different weights are put on the issue categories. There is overlap between companies scoring high on material and high on immaterial scores. Building double sort portfolios is therefore difficult. The correlation between Refinitiv scores and the immaterial score is 0.77. This is about the same as the material score. Therefore, no clear evidence that traditional ESG score contains more immaterial than material information is found.

The correlation between size and the materiality score is 0.38. Investment portfolios based on just the material score would overweigh the portfolio towards large companies. Therefore, the investment is based on the unexpected score which exhibits no relationship between the firm characteristics. The correlation between materiality and ROA of -0.04 is slightly negative. No correlations with P/B and Leverage are found which are both 0.005.

Investment portfolios end up roughly the same size because median values are used. Decile based portfolios have around 1250 observations per year, for quintile based this is 2500. Per month this is around 100 firms which is enough to construct adequately large portfolios. A portfolio of companies that have ESG data does not have the same size each year. In 2004, 879 companies have no ESG data which decreases to 98 in 2016. All analysis have 192 monthly observations, which correspond 16 years worth of data.

5. Results

5.1 Results

Table 7 and 8 present annualized alphas from high and low portfolios using a Fama and French (2015) five-factor calendar time regression. These returns have been controlled for differences in known return contributions such as size, value, profitability, and investment and can therefore be interpreted as risk adjusted returns above the market. The change column presents investment based on the change in materiality score. Differences between top and bottom equal-weighted portfolios are not found statistically significant. Value-weighted high decile portfolio has an annualized alpha of 4.00. The difference between the low portfolio of 3.65 is found moderately significant. Investment based on change is not a robust strategy because it delivers mixed results.

To test the first hypothesis, the last column in table 7 presents investment based on the unexpected change in materiality score. The high decile equal weighted portfolio has an annualized return of 5.66 percent. Subtracting the low portfolio yields an annualized alpha of 2.96 percent, which is found statistically significant at the 0.10 level. Differences between value-weighted portfolios are 4.97 and 3.78 percent. These results are found highly statistically significant. Consistent with Kahn et al. (2016) this suggest unexpected material improvement is a promising signal for investors who are looking to improve the signal-to-noise ratio present in traditional ESG scores.

Table 8 presents the differences of immaterial portfolios which tests the second hypothesis. Only the differences in quintile-based portfolios of 2.04 and 3.07 percent are found weakly statistically significant. Consistent with Kahn et al. (2016), the unexpected change of immaterial score has no influence on financial performance. Investment in firms that focus on immaterial issues does not consistency predict future stock returns.

Table 7. High low and difference material portfolio returns

			Change		Unexpected change	
			Low	High	Low	High
Equal Weighted	Decile	Annualized Alpha	3.28%	5.16%	2.70%	5.66%
		Difference		1.89%		2.96%*
	Quintile	Annualized Alpha	2.40%	3.71%	1.25%	4.10%
		Difference		1.32%		2.85%**
Value Weighted	Decile	Annualized Alpha	0.35%	4.00%	-0.63%	4.35%
		Difference		3.65%**		4.97%***
	Quintile	Annualized Alpha	0.12%	3.26%	-0.44%	3.34%
		Difference		3.14%***		3.78%***

Table 8 High low and difference immaterial portfolios

			Change		Unexpected change	
			Low	High	Low	High
Equal Weighted	Decile	Annualized Alpha	1.36%	2.45%	2.14%	2.30%
		Difference		1.09%		0.17%
	Quintile	Annualized Alpha	1.25%	3.24%	1.40%	3.38%
		Difference		1.99%*		2.02%*
Value Weighted	Decile	Annualized Alpha	0.96%	3.11%	1.64%	2.71%
		Difference		2.16%		1.08%
	Quintile	Annualized Alpha	-0.12%	2.97%	0.20%	3.30%
		Difference		3.09%*		3.07%*

***, **, * Indicate two-tailed paired p-values less than 1 percent, 2.5 percent, and 5 percent from a two-tailed paired t-test, respectively. Table 7 and 8 report alphas, factor coefficients and t-statistics from a monthly calendar-time Fama and French five-factor model regression from April 2004 to March 2020. Abnormal returns from the high portfolio minus the low portfolios are tested against the alternative that no abnormal returns are present. Annualized alpha is: $(1 + r_{montly})^{12} - 1$. Difference in alpha is the difference between the high and low portfolio. The tables report abnormal returns from portfolios based on firms scoring in the bottom (Low) and top (High) section of the material and immaterial ESG investment signal. Appendix 5 presents full tables with all factor coefficients.

Both low and high portfolios have positive alphas. Heijningen (2019) finds similar U-shaped distribution in annualized returns. Investors believe ESG deterioration is only associated with higher risk in the long run. Consistent with other literature, equal-weighted portfolios have higher annualized returns than value-weighted portfolios. Increases in rebalancing frequency and higher volatility are the main drivers for these returns. Compared to their low counterpart, high portfolios have more exposure to the CMA investment factor. Companies that improve their score invest more than companies decreasing these scores. Appendix 5 presents the same tables with all Fama and French (2015) coefficients.

Table 9. Robustness Tests Investments in Material Sustainability Issues

	Equal-Weighted			Value-Weighted		
	Low Decile	High Decile		Low Decile	High Decile	
	Annualized Alpha		Diff	Annualized Alpha		Diff
Alternative factor models						
Raw return	9.84%	12.81%	2.98%	6.31%	11.42%	5.11%***
Three-factor	3.24%	6.28%	3.04%*	0.07%	4.89%	4.82%***
Four-factor with momentum	3.98%	6.69%	2.71%*	0.40%	5.18%	4.78%***
Alternative weights						
Unweighted materiality matrix	3.67%	4.91%	1.24%	0.45%	4.07%	3.63%**
Subset of firms						
Excluding Human capital firms	2.46%	5.69%	3.23%	-1.02%	3.68%	4.7%**
Excluding polluting firms	1.85%	6.15%	4.3%**	-0.52%	5.28%	5.8%***
Excluding below mean small firms	1.93%	5.52%	3.58%**	-1.46%	4.39%	5.85%***
Subset of periods						
2004 –2011	5.20%	7.90%	2.71%	1.41%	5.70%	4.28%**
2012 –2019	-3.87%	1.09%	4.96%***	-6.33%	0.51%	6.83%***

***, **, * Indicate two-tailed paired p-values less than 1 percent, 2.5 percent, and 5 percent from a two-tailed paired t-test, respectively. Table 9 reports alphas from Fama and French (2015) five-factor calendar time regressions of monthly returns for the period April 2004 – March 2020. Abnormal returns from the high portfolio minus the low portfolios are tested against the alternative that no abnormal returns are present. Under "Alternative models" differences in raw returns as well as the Fama French (1993) 3-factor model and Carhart (1997) 4-factor model are reported. Under "Alternative weights" the results using a materiality matrix that is not weighted by the number of industries are reported. Under "Subset of firms" the sample excluding several SASB sectors which relate to human capital intensive or polluting firms, also separately for only large firms are reported. Under "Sub periods" reports the regression separately for the period from 2004–2011 and 2012–2020. Appendix 6 presents full tables with all factor coefficients.

To test the validity of the model, tests are performed on the residuals. A Breusch and Pagan p-value of 0.34 indicates there is no heteroskedasticity in the standard errors. A Durbin and Watson autocorrelation test yields d-statistic value of 1.79. This indicates a slight positive autocorrelation, which is expected in financial data. Returns are tailed in the positive direction, indicated by a kurtosis value of 5. An ARCH coefficient of 0.22 and GARCH coefficient of 0.70 from a GARCH (1,1) model. The conditional variance is stationary as both coefficients

together are below 1. The model predicts a nonnegative variance as both coefficients are positive.

Table 9 presents a robustness analyses using alternative factor models, weights, and subsets of the sample for material equal-weighted decile portfolio. Robustness tests based on immaterial signals are presented in appendix 5, since immaterial portfolios are mostly found insignificant in table 8. Raw returns (i.e., no risk adjustment) show an outperformance of 5.11 and 2.98 percent in value- and equal-weighted portfolios, respectively. In a Carhart (1997) four-factor model outperformance of low portfolio is 4.78 percent for value-weighted and 2.71 for equal weighted. The purposed ESG change investment strategy is more than a proxy for a price momentum strategy.

Using an unweighted matrix only value weighted portfolio outperformance of 3.62 is found weakly statistically significant. Because more noise is added to the signal, this result is expected. Exclusion of polluting industries does not produce smaller differences. For value and equal-weighted portfolios the difference is increased to 5.80 and 4.30 percent respectively Less pronounced results where found by excluding human capital industries where only the value-weighted portfolios outperform by 4.96 percent.

Table 10. Difference between high material, immaterial and Refinitiv ESG score.

			Material ESG issues	Refinitiv ESG	Immaterial ESG issues
Equal Weighted	Decile	Annualized Alpha	5.66%	3.19%	2.31%
		Difference		-2.46%	-3.35%***
	Quintile	Annualized Alpha	4.10%	3.95%	3.38%
		Difference		-0.152%	-0.72%
Value Weighted	Decile	Annualized Alpha	4.35%	4.67%	2.71%
		Difference		0.32%	-1.63%*
	Quintile	Annualized Alpha	3.38%	4.12%	3.29%
		Difference		0.20%	-0.04%

***, **, * Indicate two-tailed paired p-values less than 1 percent, 2.5 percent, and 5 percent from a two-tailed paired t-test, respectively. Table 10 reports alphas from Fama and French (2015) five-factor calendar time regressions of monthly returns for the period April 2004 – March 2020. Abnormal returns from the high portfolio minus the low portfolios are tested against the alternative that no abnormal returns are present. Annualized alpha is found by $(1 + r_{montly})^{12} - 1$. Difference in alpha is the difference between the high and low portfolio. The tables report results for portfolios from firms scoring in the top (High) section of the material, immaterial and Refinitiv ESG investment signal.

From 2012-2020 differences appear to be stronger. For equal-weighted differences are 6.38 percent and value weighted 4.96 percent. In this period investor valued sustainability investments higher and found ESG topics more important. Large differences in performance

are found between equal and value weighted investment strategies. In Kahn et al (2016) equal decile weighted portfolios outperform by 2.69 percent and value-weighted by 7.47 percent.

Table 10 directly compares material to immaterial and normal ESG investments, consistent with the third hypothesis. Annualized alphas from high material ESG issues where already found in table 7. Only the differences between immaterial portfolios of 3.35 and 1.63 percent are found statistically significant. These difference can not be arbitraged by a short, long strategy because identical stocks can be present in both portfolios. Because quintile immaterial portfolio differences are not found significant the results remain mixed.

Material ESG scores do not significantly improve the noise which is present in traditional ESG scores. An explanation for these results is within the score calculation of Refinitiv, there already exists a form material classification. The dynamic magnitude matrix utilized is based on relative disclosure of issues in an industry. However, traditional ESG research rarely finds a positive link between financial performance. Therefore, the distinguishment between material and immaterial ESG issues is still important.

Table 11. Performance on double sort portfolios

			HMLI	HMHI	LMHI	LMLI
Equal Weighted	Decile	Annualized Alpha	3.26%	3.52%	2.58%	2.57%
		Difference		0.25%	-0.68%	-0.69%
	Quintile	Annualized Alpha	2.40%	2.23%	2.92%	3.51%
		Difference		0.83%	0.53%	1.11%
Value Weighted	Decile	Annualized Alpha	2.04%	3.64%	2.01%	0.16%
		Difference		1.60%	-0.04	-1.80%*
	Quintile	Annualized Alpha	1.08%	2.50%	0.91%	0.95%
		Difference		1.42%	-0.17%	-0.13%

***, **, * Indicate two-tailed paired p-values less than 1 percent, 2.5 percent, and 5 percent from a two-tailed paired t-test, respectively. Table 11 reports alphas from Fama and French (2015) five-factor calendar time regressions of monthly returns for the period April 2004 – March 2020. Abnormal returns from the HMLI portfolio minus the other portfolios are tested against the alternative that no abnormal returns are present. Annualized alpha is: $(1 + r_{montly})^{12} - 1$. Difference in alpha is the difference between the high and low portfolio. The tables report results for portfolios from firms scoring in the bottom (Low) and top (High) section of the material and immaterial ESG investment signal. Double sorted scores are found using the following formulas: $HMLI_{i,t} = Material\ score_{i,t,normalized} - Immaterial\ score_{i,t,normalized}$,$HMHI_{i,t} = Material\ score_{i,t,normalized} + Immaterial\ score_{i,t,normalized}$, $LMLI_{i,t} = -Material\ score_{i,t,normalized} + Immaterial\ score_{i,t,normalized}$, $LMLI_{i,t} = -Material\ score_{i,t,normalized} - Immaterial\ score_{i,t,normalized}$

Forming double sort portfolios based on high materiality and low immateriality (HMLI) issues is difficult because firms ranking high on material issues also rank high on immaterial issues. Low materiality portfolios in the last two columns in table 11 have mostly negative differences.

Only the difference between the LMLI of 1.8 percent is found weakly statistically significant. Kahn et al. (2016) finds the HMLI portfolio outperform HMHI, LMHI and LMLI portfolios, this paper does not replicate these results. Contrarily to these findings, the HMHI portfolios appear to have a higher abnormal returns than HMLI portfolios, but these results are not found statistically significant.

Table 12 shows investment in portfolios with and without ESG data available. Especially in the early years of the sample most firms have no ESG data available. Making this distinction is often overlooked in ESG research. The exclusion of these companies from the start would lead to a survivorship bias as the scores are often determined ex-post. The difference of 8.49 percent for equal-weighted portfolios is found highly statistically significant. However, another plausible explanation is that companies without data have only been recently added to the index. Companies added recently have high abnormal returns since these just made it to a large market capitalization index.

Table 12. Comparison performance firms with and without ESG data

		With ESG data	Without ESG data
Equal Weighted	Annualized Alpha	2.67%	11.12%
	Difference		8.49%***
Value Weighted	Annualized Alpha	0.47%	12.23%
	Difference		11.76%***

***, **, * Indicate two-tailed paired p-values less than 1 percent, 2.5 percent, and 5 percent from a two-tailed paired t-test, respectively. Table 12 reports alphas from a Fama and French (2015) five-factor calendar time regressions from monthly returns for the period April 2004 – March 2014. Abnormal returns from the without ESG portfolio minus the with ESG portfolio is tested against the alternative that no abnormal returns are present. Annualized alpha is: $(1 + r_{montly})^{12} - 1$. Difference is alpha is the difference between the portfolios. Without ESG data are companies which have no ESG data available but are included in the market index at that time. With ESG data are all companies which do have data available and are used in the rest of the study.

5.2 Discussion on using unexpected change

In addition to using change, Kahn et al. (2016) propose to use unexpected change as the investment input. Rather than investing in material ESG scores, the paper invests in unexpected change of material ESG scores. This paper takes more caution with unexpected change and includes change, because it is more grounded in ESG literature. Unexpected change leads in more cases to significant financial outperformance than only using change.

Unexpected score level is compelling to compensate for the size bias present in ESG scores (Drempetic et al., 2020). Unexpected score change based on growth opportunities, valuation, financial structure, and investment profile is not used in other literature. The relationship between change in size and change in score is also not grounded in literature. Table 7 shows a

negative significant coefficient of -0.71 between the change in size and change in ESG score. Therefore, growing firms are expected to decrease their score. This does not rhyme with the assumption that larger firms have higher ESG scores. Furthermore, companies that have increased in size need smaller improvements in material ESG score to be included in the top portfolio.

Explanatory variables were chosen because of their high correlation with score level. However, these variables do not have any correlation with score change. Between size level the correlation is 0.38 where for size change this is 0.01. Using the unexpected score adds size, valuation, and capital expenditure factors into the investment signal. These factors are later controlled for in Fama and French (2015) five-factor regression model. Further investigation in the relationship between unexpected change and known factors that contribute to returns is necessary. Using unexpected level can be explained by intuition and literature. For unexpected change both are not possible.

6. Conclusion

A similar methodology to Kahn et al. (2016) is used to identify material ESG topics. Besides actual performance improvements this paper investigated the disclosure of policies that mitigate material issues. Using a completely different sustainability dataset, the main conclusions remain the same. Firms that unexpectedly improve material ESG scores financially outperform firms scoring inferior on these scores. Therefore, material ESG disclosure can improve the signal-to-noise ratio present in traditional sustainability investing. Unexpected improvements isolate companies that are superior in managing material sustainability risks. Alternatively, immaterial ESG score improvements do not enhance financial performance.

Compared to traditional ESG, the SASB identifies material ESG issues which are likely to impact financial conditions or operating performance. For institutional investors committed to ESG initiatives, material scores are more important to stock performance than immaterial scores. Therefore, materiality is proven to be a valuable distinction to make within ESG investing. Companies can benefit from strategic ESG investments into industry-specific material ESG issues. Within the field of ESG investing, materiality is significant and should be considered by asset managers, institutional investors, and companies.

Companies and ESG rating agencies are beginning to recognize the differences between material and immaterial issues. Thomson Reuters overweighs material ESG issues in the

Refinitiv ESG score. Presumably because of this, material issues do not outperform Refinitiv ESG scores. Using the Thomson Reuters dataset, forming double sort portfolios is difficult due to a high correlation between material and immaterial scores. Therefore, companies that improve material and decrease immaterial scores do not financially outperform other variations of double sort portfolios.

A ESG change strategy evaluates companies which improve on sustainability issues. Compared to companies with high ESG scores, these firms might be far from adequately managing their sustainability risks. Investment strategies using score change is sometimes seen as not being sustainable investments, but only having short-term benefits. Alphas from this study are based on a hypothetical portfolio with no limits on turnover and yearly rebalancing. Furthermore, transaction cost and other short selling constrains are not taken into consideration.

Different material topics arise within value chains, between companies or across geographical locations. Therefore, to further understand how materiality adds value, future research should evaluate the SASB framework more extensively in specific industries. Cross-sectional and within firm-level panel regression would further verify the conclusions. In material research, unexpected change scores are found using a methodology build to find unexpected levels. The interior workings of the unexpected change component need to be further understood.

Materiality is an important distinguishment to make in ESG investing. Companies should focus disclosure on issues which are important to its audiences. The SASB framework offers an important guide to companies wanting to improve material topics relevant to business performance. Investors can use the framework to engage businesses not adequately responding to material ESG risks and opportunities. Fortunately, the importance of material ESG topics is starting to receive more recognition by companies, investors, rating agencies and researchers.

7. References

Adams, C. A. (2017). Conceptualizing the contemporary corporate value creation process. Accounting, Auditing & Accountability Journal, 30(4), 906–931.

Berg, F., Kölbel, J.F., & Rigobón, R. (2020). Aggregate Confusion: The Divergence of ESG Ratings.

Blitz, D., Fabozzi, Frank J. (2017). Sin Stocks Revisited: Resolving the Sin Stock Anomaly. Journal of Portfolio Management, Vol. 44, No. 1, 2017

Brammer, S., & Millington, A. (2008). Does it pay to be different? An analysis of the relationship between corporate social and financial performance. Strategic Management Journal, 29(12), 1325–1343.

Brown, H. S., de Jong, M., & Lessidrenska, T. (2009). The rise of the global reporting initiative: a case of institutional entrepreneurship. Environmental Politics, 18(2), 182–200.

Carhart, M. M. (1997). On persistence in mutual fund performance. The Journal of finance, 52(1), 57-82.

CFA Institute (2017). Global Perceptions of Environmental, Social, and Governance Issues in Investing. Retrieved from https://www.cfainstitute.org/en/research/survey-reports/esg-survey-2017

Derwall, J., Koedijk, K., & Ter Horst, J. (2011). A tale of values-driven and profit-seeking social investors. Journal of Banking & Finance, 35(8), 2137-2147.

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The eco-efficiency premium puzzle. Financial Analysts Journal, 61(2), 51-63.

Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: corporate sustainability ratings under review. Journal of Business Ethics, 167(2), 333–360.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835-2857.

European parliament and the council (2019). 2019/2088 on sustainability-related disclosures in the financial services sector. Official Journal of the European Union, L 317/1

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. (1993). Common risk factors in the returns on stocks and bonds. Journal of.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210-233.

Brown, H. S., de Jong, M., & Levy, D. L. (2009). Building institutions based on information disclosure: lessons from GRI's sustainability reporting. Journal of cleaner production, 17(6), 571-580.

Heijningen, K. (2019). The impact of ESG factor materiality on stock performance of firms. Unpublished working paper. Rotterdam School of management Erasmus University, Netherlands.

Henisz, W. J., & McGlinch, J. (2019). ESG, material credit events, and credit risk. Journal of Applied Corporate Finance, 31(2), 105–117.

Henriksson et al. (2019). SASB Focus Helps Ensure ESG Practices Are Rewarding. SASB ESG Integration Insights, 7- 11.

Jensen, M. C. (2010). Value maximization, stakeholder theory, and the corporate objective function. Journal of Applied Corporate Finance, 22(1), 32–42.

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: first evidence on materiality. Accounting Review, 91(6), 1697–1724.

Kotsantonis, S., Bufalari, V. (2019). Do sustainable banks outperform? Driving value creation through ESG practices

Lo, S. F., & Sheu, H. J. (2007). Is corporate sustainability a value-increasing strategy for business? Corporate Governance: An International Review, 15(2), 345-358.

Lydenberg, S., Rogers, J. & Wood, D., (2012) From Transparency to Performance: Industry-Based Sustainability Reporting on Key Issues.

Mervelskemper, L., & Streit, D. (2017). Enhancing market valuation of ESG performance: is integrated reporting keeping its promise? Business Strategy and the Environment, 26(4), 536–549.

Nagy, Z., Kassam, A., & Lee, L. E. (2016). Can ESG add alpha? An analysis of ESG tilt and momentum strategies. The Journal of Investing, 25(2), 113-124.

SASB Conceptual framework (2020) Retrieved from https://www.sasb.org/wp-content/uploads/2020/02/SASB_Conceptual-Framework_WATERMARK.pdf

Steinbarth, E. (2018) Materiality matters: Targeting the ESG issues that can impact performance. Russel investments

Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: evidence from a global survey. Financial Analysts Journal, 74(3), 87–103.

8. Appendix

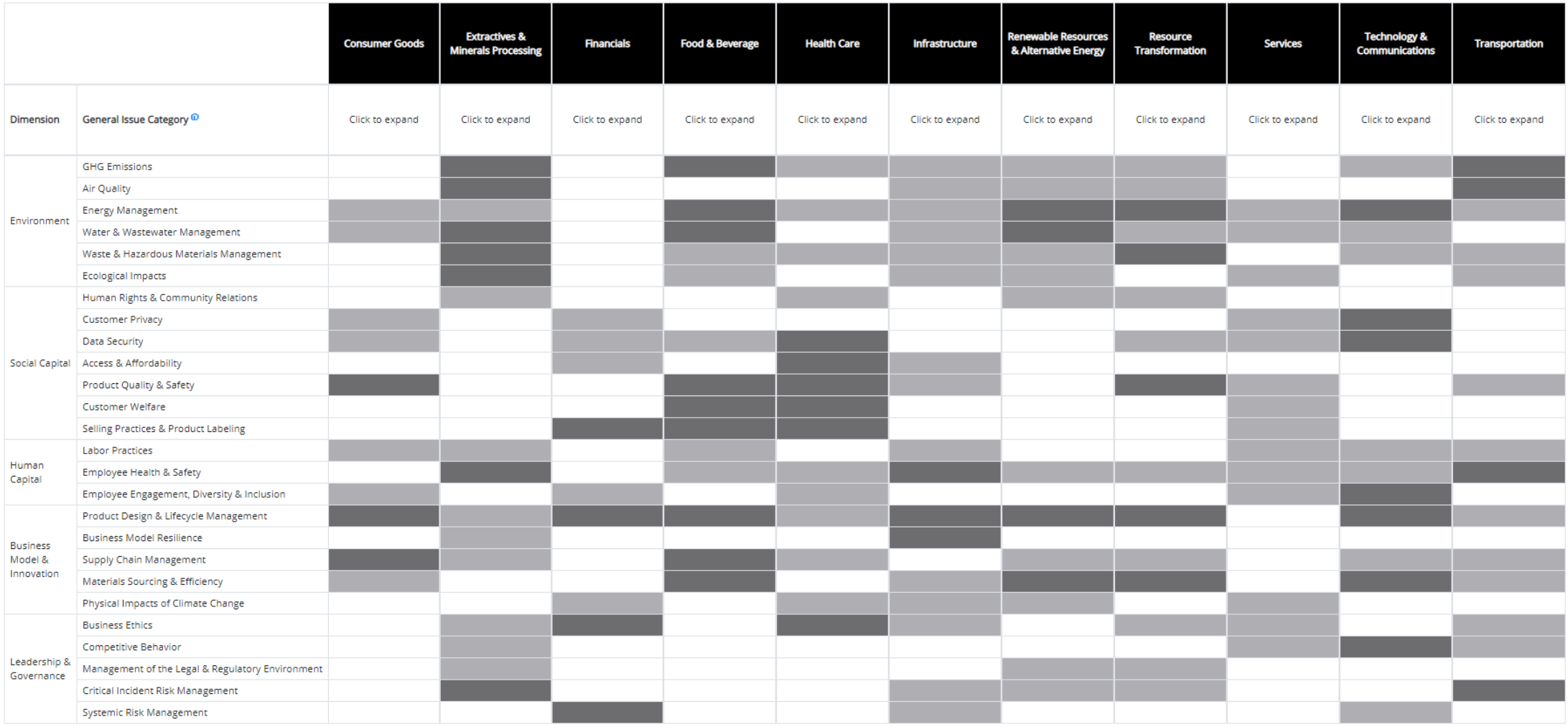
Appendix 1: Differences between ESG datasets

Although sometimes being treated this way green ratings are not the same thing as credit ratings. Where credit ratings have been around for a long time and correlate with on average 0.9 between agencies. The six biggest ESG raters, only have an average correlation 0.54 (Berg et al., 2020). In an example made by the Wall street journal (2018) Tesla only has a 53 rating in Thomson Reuters because of their lack in disclosure of carbon data. However, in MSCI database they receive a triple AAA rating because of their outstanding sustainable business model. Where sustainability is often seen as a homogenous term, it is important to understand that these are heterogeneous score which differ across rating companies.

According to Sustainalytics which is owned by Morningstar, a company is both exposed to manageable and unmanageable ESG risks. Their score reflects the amount of managed ESG risk compared to the total ESG risk faced by that company. This approach has similarities to materiality, as the overall ESG risk faced by a company varies between industries. Idiosyncratic risk ratings are incorporated in their overall score, for making the ratings more company specific. This has the downside to make the score more complex. The MSCI ESG ratings focuses in on what is significant to a company's bottom line performance and compare these with its peer group. The same as Thomson Reuters, MSCI assigns percentage weights to each ESG risk to measure a company's resilience to long-term, financially relevant scenarios. By incorporating key issues, MSCI adds issues that are material to a specific company to the eventual score. A special feature from the Thomson Reuters database is that many datapoints are utilized as a yes or no question and transform these into percentile rank score. Other datasets use score value immediately as inputs for their indicators. This two-sided answer allows for a more unbiased interpretation but has the downside that additional information is lost.

Berg et al. (2020) find many instances of measurement, scope, and weight divergence between the rating agencies. Measurement divergence between databases problematic, since is not merely driven by difference in opinions about weights, but also by disagreement about facts. The different usage of indicators undermines the legitimacy of these ratings, which should ultimately be based on facts that can be ascertained. Companies can take advantage of the discrepancies between the ratings to focus on the more favorable. Therefore, it is important to understand the differences between scores, to incorporate their characteristics in formulating your question while doing research.

Appendix 2: SASB materiality map on sector level



Dimension	General Issue Category	Consumer Goods	Extractives & Minerals Processing	Financials	Food & Beverage	Health Care	Infrastructure	Renewable Resources & Alternative Energy	Resource Transformation	Services	Technology & Communications	Transportation
		Click to expand	Click to expand	Click to expand	Click to expand	Click to expand	Click to expand	Click to expand	Click to expand	Click to expand	Click to expand	Click to expand
Environment	GHG Emissions											
	Air Quality											
	Energy Management											
	Water & Wastewater Management											
	Waste & Hazardous Materials Management											
	Ecological Impacts											
Social Capital	Human Rights & Community Relations											
	Customer Privacy											
	Data Security											
	Access & Affordability											
	Product Quality & Safety											
	Customer Welfare											
	Selling Practices & Product Labeling											
Human Capital	Labor Practices											
	Employee Health & Safety											
	Employee Engagement, Diversity & Inclusion											
Business Model & Innovation	Product Design & Lifecycle Management											
	Business Model Resilience											
	Supply Chain Management											
	Materials Sourcing & Efficiency											
	Physical Impacts of Climate Change											
Leadership & Governance	Business Ethics											
	Competitive Behavior											
	Management of the Legal & Regulatory Environment											
	Critical Incident Risk Management											
	Systemic Risk Management											

Appendix 3: Mapping of ESG datapoints to SASB issue categories

	Consumer Goods	Extractives & Miniral processing	Financials	Food and Beverage	Health Care	Infrastructure	Renewable Recources & Alternative Energy	Recource transformation	Services	Technology & Communication	Transportation
GHG Emissions		TargetEmissions		TargetEmissions	TargetEmissions	TargetEmissions	TargetEmissions	TargetEmissions		TargetEmissions	TargetEmissions
Air Quality		PolicyEmissions				PolicyEmissions	PolicyEmissions	PolicyEmissions			PolicyEmissions
Energy Management	PolicyEnergyEfficiency	PolicyEnergyEfficiency		GreenBuildings	GreenBuildings	GreenBuildings	PolicyEnergyEfficiency	PolicyEnergyEfficiency	GreenBuildings	GreenBuildings	PolicyEnergyEfficiency
Water & Wastewater Management	PolicyWaterEfficiency	PolicyWaterEfficiency		PolicyWaterEfficiency		PolicyWaterEfficiency	PolicyWaterEfficiency	PolicyWaterEfficiency	PolicyWaterEfficiency	PolicyWaterEfficiency	
Waste & Hazardous Materials Management		ToxicChemicalsReduction		WastereductionInitiatives	ToxicChemicalsReduction	WastereductionInitiatives	ToxicChemicalsReduction	ToxicChemicalsReduction		eWasteReduction	WastereductionInitiatives
Ecological Impacts		BiodiversityImpactReduction		BiodiversityImpactReduction		BiodiversityImpactReduction	BiodiversityImpactReduction		BiodiversityImpactReduction		BiodiversityImpactReduction
Human Rights & Community Relations		ILOFundamentalHumanRights			PolicyHumanRights		PolicyChildLabor	ILOFundamentalHumanRights			
Customer Privacy	PolicyDataPrivacy		PolicyDataPrivacy						PolicyDataPrivacy	PolicyDataPrivacy	
Data Security	PolicyCybersecurity		PolicyCybersecurity	PolicyDataPrivacy	PolicyDataPrivacy			PolicyCybersecurity	PolicyDataPrivacy	PolicyCybersecurity	
Access & Affordability			ProductAccessLowPrice		ProductAccessLowPrice	ProductAccessLowPrice					
Product Quality & Safety	ProductQualityMonitoring			ProductQualityMonitoring	ProductQualityMonitoring	ProductQualityMonitoring		ProductQualityMonitoring	ProductQualityMonitoring		ProductQualityMonitoring
Customer Welfare				PolicyCustomerHealthSafety	PolicyCustomerHealthSafety				PolicyCustomerHealthSafety		
Selling Practices & Product Labeling			Policyresponsiblemarketing	PolicySustainablePackaging	Policyresponsiblemarketing				Policyresponsiblemarketing		
Labor Practices		AnalyticTrainingDevPolicy		AnalyticTrainingDevPolicy		AnalyticTrainingDevPolicy			AnalyticTrainingDevPolicy	AnalyticTrainingDevPolicy	AnalyticTrainingDevPolicy
Employee Health & Safety		EmployeesHealthSafetyTeam		EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam	EmployeesHealthSafetyTeam
Employee Engagement, Diversity & Inclusion	EmployeeresourceGroups		EmployeeresourceGroups		EmployeeresourceGroups				EmployeeresourceGroups	EmployeeresourceGroups	
Product Design & Lifecycle Management	PolicySustainablePackaging	AnalyticProductImpactMin	EnvAUM	PolicySustainablePackaging	AnalyticProductImpactMin	AnalyticProductImpactMin	AnalyticProductImpactMin	AnalyticProductImpactMin		Takebackrecyclinginitiatives	AnalyticProductImpactMin
Business Model Resilience		CSRreportingglobalactivities				CSRreportingglobalactivities					
Supply Chain Management	Policyenvsupplychain	Policyenvsupplychain		Policyenvsupplychain	Policyenvsupplychain		Policyenvsupplychain	Policyenvsupplychain		Policyenvsupplychain	Policyenvsupplychain
Materials Sourcing & Efficiency	analyticresourceredpolicy			analyticresourceredpolicy		EnvMaterialsSourcing	EnvMaterialsSourcing	EnvMaterialsSourcing		analyticresourceredpolicy	EnvMaterialsSourcing
Physical Impacts of Climate Change			CSRreporting		CSRreporting	CSRreporting	CSRreporting		CSRreporting		
Business Ethics		PolicyBusinessEthics	PolicyBusinessEthics		PolicyBusinessEthics	PolicyBusinessEthics		PolicyBusinessEthics	PolicyBusinessEthics		PolicyBusinessEthics
Competitive Behavior		PolicyFairCompetition							PolicyFairCompetition	PolicyFairCompetition	PolicyFairCompetition
Management of the Legal & Regulatory Environment		CSRReportingexternalaudit					CSRReportingexternalaudit	CSRReportingexternalaudit			
Critical Incident Risk Management		Crisismgtsystems				Crisismgtsystems	Crisismgtsystems	Crisismgtsystems			Crisismgtsystems
Systemic Risk Management			envprojectFinancing			SustainabilityComm				envprojectFinancing	

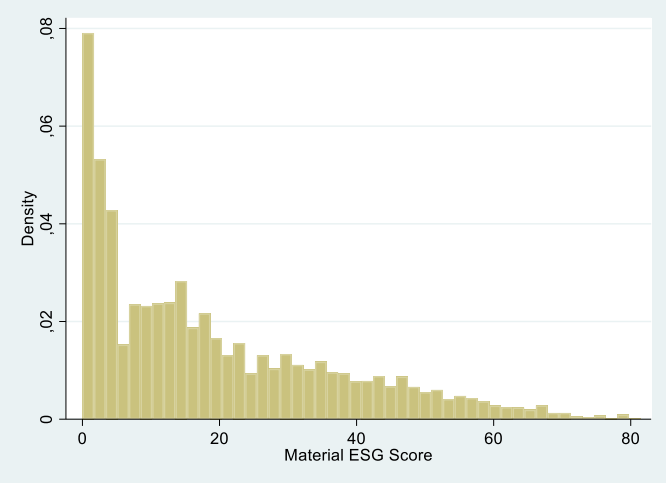
Appendix 4: Distribution of material and Refinitiv ESG score

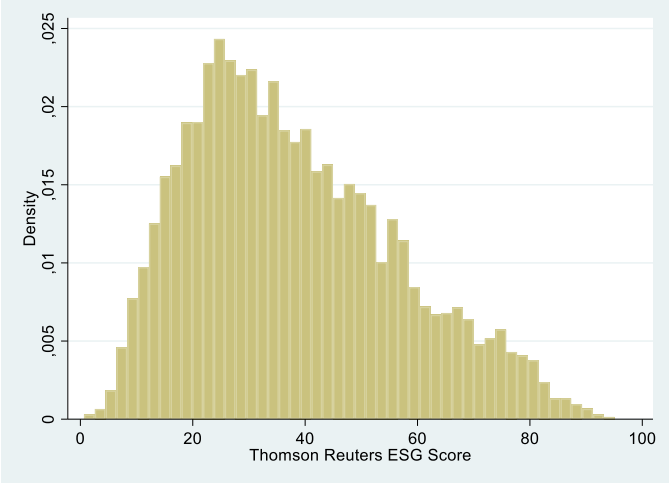
Material ESG scores are more skewed towards lower numbers because of the implementation rate of the issues used to construct the score. Only in the last five years have companies started to write policies on how to deal with data privacy. This issue has a larger share in material score compare to the Thomson Reuters ESG score.

In the second period, the scores have a more similar distribution. Companies have improved material ESG practices more than all ESG practices. Additionally, the score is less skewed towards lower numbers.

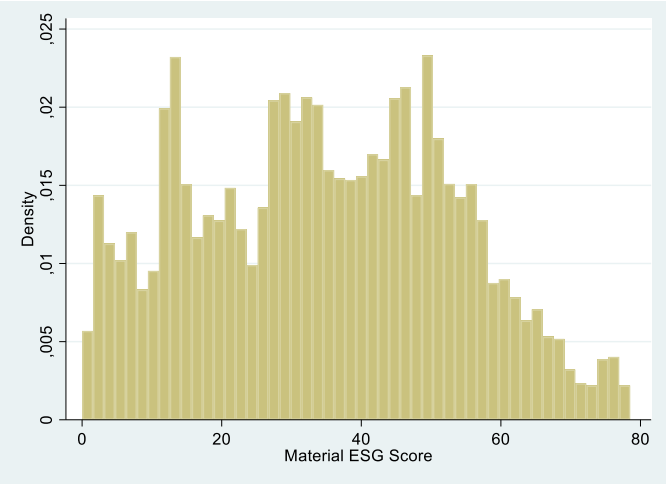
Histogram of scores which are used as the input for investments

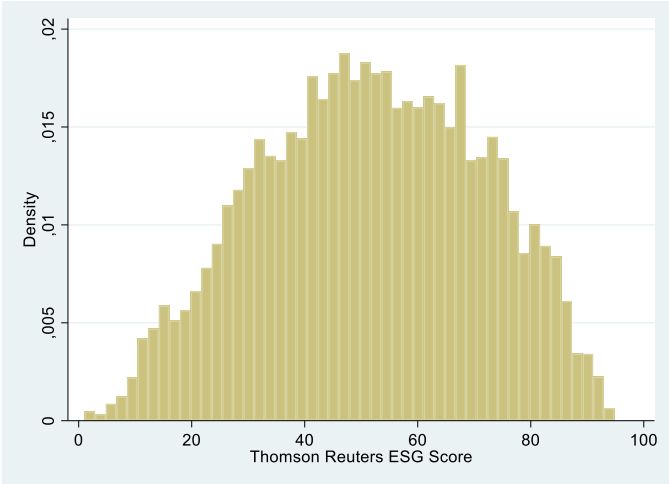
Period 2003 - 2012





Period 2012 - 2020





Appendix 5:Tables main results with Fama and French factors

Table 7.1. Calendar Time portfolio analysis equal-weighted decile material portfolios

	Change in materiality				Unexpected change in materiality			
	LOW Decile		HIGH Decile		LOW Decile		HIGH Decile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0 .0027	1 .42	0 .0042	2 .76	0 .0022	1 .09	0 .0046	3 .14
Mkt_RF	0 .9423	18 .06	0 .8783	20 .93	0 .9521	16 .92	0 .8768	21 .69
SMB	0 .1742	2 .00	0 .1477	2 .11	0 .1674	1 .79	0 .1646	2 .45
HML	0 .0687	0 .83	-0 .0309	-0 .47	0 .1285	1 .45	-0 .0578	-0 .91
RMW	0 .151	1 .02	0 .2281	1 .92	0 .1101	0 .69	0 .2531	2 .22
CMA	0 .1724	1 .36	0 .1315	1 .29	0 .1163	0 .85	0 .1074	1 .10
Observations	192		192		192		192	
R-squared	0 .73		0 .78		0 .71		0 .79	
Annualized Alpha	3 .28%		5 .16%		2 .7%		5 .66%	
Difference in Alpha			1 .88%				2 .96%*	

Table 8.1. Calendar Time portfolio analysis equal-weighted decile immaterial portfolios

	Change in immateriality				Unexpected change immateriality			
	LOW Decile		HIGH Decile		LOW Decile		HIGH Decile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0 .0011	0 .77	0 .002	1 .39	0 .0018	1 .06	0 .0019	1 .32
Mkt_RF	0 .9492	23 .54	0 .9311	23 .14	0 .9347	20 .28	0 .9315	23 .41
SMB	0 .1049	1 .56	0 .0939	1 .40	0 .1316	1 .71	0 .0891	1 .34
HML	0 .0328	0 .52	0 .0224	0 .35	0 .0704	0 .97	-0 .0012	-0 .02
RMW	0 .1349	1 .18	0 .1719	1 .51	0 .1593	1 .22	0 .1888	1 .68
CMA	0 .2424	2 .48	0 .1746	1 .79	0 .2181	1 .95	0 .1547	1 .61
Observations	192		192		192		192	
R-squared	0 .81		0 .81		0 .77		0 .81	
Annualized Alpha	1 .36%		2 .45%		2 .14%		2 .31%	
Difference in Alpha			1 .09%				0 .17%	

Table 7.2. Calendar Time portfolio analysis equal-weighted quintile material portfolios

	Change in materiality				Unexpected change in materiality			
	LOW Quintile		HIGH Quintile		LOW Quintile		HIGH Quintile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.002	1.28	0.003	2.02	0.001	0.57	0.0034	2.25
Mkt_RF	0.9415	22.24	0.9161	22.03	0.9703	19.35	0.9194	22.33
SMB	0.1435	2.03	0.165	2.38	0.1775	2.13	0.1487	2.17
HML	0.1279	1.91	-0.0248	-0.38	0.1519	1.92	-0.037	-0.57
RMW	0.1104	0.92	0.215	1.83	0.1674	1.18	0.2136	1.84
CMA	0.1393	1.36	0.148	1.47	0.1315	1.08	0.1442	1.45
Observations	192		192		192		192	
R-squared	0.81		0.79		0.77		0.8	
Annualized Alpha	2.39%		3.71%		1.25%		4.1%	
Difference in Alpha			1.32%				2.85%**	

Table 8.2. Calendar Time portfolio analysis equal-weighted quintile immaterial portfolios

	Change in immateriality				Unexpected change immateriality			
	LOW Quintile		HIGH Quintile		LOW Quintile		HIGH Quintile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.001	0.72	0.0027	1.87	0.0011	0.7	0.0028	1.95
Mkt_RF	0.963	24.26	0.9351	23.84	0.9794	22.32	0.9424	23.98
SMB	0.1063	1.61	0.1107	1.69	0.1568	2.15	0.1069	1.63
HML	0.0071	0.11	0.056	0.9	0.0878	1.27	0.0525	0.85
RMW	0.2251	2.01	0.0902	0.81	0.2219	1.79	0.0955	0.86
CMA	0.1739	1.81	0.214	2.25	0.166	1.56	0.2144	2.25
Observations	192		192		192		192	
R-squared	0.82		0.82		0.8		0.82	
Annualized Alpha	1.25%		3.24%		1.34%		3.38%	
Difference in Alpha			1.99%*				2.04%**	

Table 7.3. Calendar Time portfolio analysis value-weighted decile material portfolios

	Change in materiality				Unexpected change in materiality			
	LOW Quintile		HIGH Quintile		LOW Quintile		HIGH Quintile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0003	0.19	0.0033	2.5	-0.0005	-0.35	0.0036	2.71
Mkt_RF	0.916	21.76	0.9008	25.01	0.9157	22.39	0.8827	24.39
SMB	-0.096	-1.37	-0.0353	-0.59	-0.1474	-2.16	-0.0188	-0.31
HML	0.0216	0.33	0.0008	0.01	0.0636	0.99	-0.0071	-0.12
RMW	0.1627	1.37	0.1315	1.29	0.1922	1.66	0.1178	1.15
CMA	0.2134	2.09	0.1199	1.37	0.1432	1.45	0.1115	1.27
Observations	192		192		192		192	
R-squared	0.76		0.82		0.78		0.81	
Annualized Alpha	0.35%		4.00%		-0.63%		4.35%	
Difference in Alpha			3.65%**				4.97%***	

Table 8.3. Calendar Time portfolio analysis value-weighted decile immaterial portfolios

	Change in immateriality				Unexpected change immateriality			
	LOW Quintile		HIGH Quintile		LOW Quintile		HIGH Quintile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0008	0.63	0.0026	2.07	0.0014	1.07	0.0022	1.78
Mkt_RF	0.8905	25.48	0.9103	26.67	0.878	25.13	0.913	26.42
SMB	-0.0755	-1.3	-0.0765	-1.35	-0.1051	-1.81	-0.0561	-0.97
HML	-0.025	-0.45	-0.1348	-2.5	-0.0083	-0.15	-0.1495	-2.74
RMW	0.07	0.71	0.2331	2.42	0.074	0.75	0.2372	2.43
CMA	0.3695	4.36	0.2669	3.23	0.339	4.01	0.2835	3.39
Observations	192		192		192		192	
R-squared	0.81		0.82		0.81		0.82	
Annualized Alpha	0.96%		3.11%		1.64%		2.71%	
Difference in Alpha			2.16%				1.08%	

Table 7.4. Calendar Time portfolio analysis value-weighted quintile material portfolios

	Change in materiality				Unexpected change in materiality			
	LOW Quintile		HIGH Quintile		LOW Quintile		HIGH Quintile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0001	0.08	0.0027	2.15	-0.0004	-0.27	0,0027	2,19
Mkt_RF	0.9375	27.57	0.9137	26.64	0.9519	25.79	0,9296	26,9
SMB	-0.1226	-2.17	-0.0664	-1.16	-0.141	-2.29	-0,0826	-1,44
HML	0.1208	2.25	0.0107	0.2	0.115	1.97	-0,0088	-0,16
RMW	0.1566	1.63	0.0048	0.05	0.1837	1.76	0,0272	0,28
CMA	0.1688	2.05	0.1791	2.15	0.1586	1.77	0,1926	2,3
Observations	192		192		192		192	
R-squared	0.85		0.83		0.83		0,83	
Annualized Alpha	0.12%		3.26%		-0.44%		3,34%	
Difference in Alpha			3.14%***				3.78%***	

Table 8.4. Calendar Time portfolio analysis value-weighted quintile immaterial portfolios

	Change in immateriality				Unexpected change immateriality			
	LOW Quintile		HIGH Quintile		LOW Quintile		HIGH Quintile	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	-0.0001	-0.08	0.0024	2.11	0.0002	0.15	0,0027	2,35
Mkt_RF	0.9083	26.37	0.954	29.85	0.9225	25.73	0,9596	30,17
SMB	-0.1131	-1.97	-0.114	-2.14	-0.1049	-1.76	-0,1224	-2,31
HML	-0.0362	-0.67	0.0367	0.73	0.0143	0.25	0,0351	0,7
RMW	0.1487	1.53	0.0035	0.04	0.1767	1.74	0,0097	0,11
CMA	0.3323	3.98	0.1917	2.48	0.3726	4.29	0,1835	2,38
Observations	192		192		192		192	
R-squared	0.82		0.86		0.81		0,86	
Annualized Alpha	-0.12%		2.97%		0.23%		3,3%	
Difference in Alpha			3.09%*				3.07%*	

Table 10.1 Equal-weighted decile difference between material, immaterial and Thomson Reuters ESG score.

	Material		Thomson Reuters		Immaterial	
	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0046	3.14	0.0026	1.7	0.0019	1.32
Mkt_RF	0.8768	21.69	0.9591	22.48	0.9315	23.41
SMB	0.1646	2.45	0.1817	2.56	0.0891	1.34
HML	-0.0578	-0.91	0.0223	0.33	-0.0012	-0.02
RMW	0.2531	2.22	0.065	0.54	0.1888	1.68
CMA	0.1074	1.1	0.2158	2.09	0.1547	1.61
Observations	192		192		192	
R-squared	0.79		0.81		0.81	
Annualized Alpha	5.66%		3.19%		2.31%	
Difference in Alpha			-2.46%		-3.35%***	

Table 10.2 Equal-weighted quintile difference between material. immaterial and Thomson Reuters ESG score

	Material		Thomson Reuters		Immaterial	
	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0034	2.25	0.0032	2.17	0.0028	1.95
Mkt_RF	0.9194	22.33	0.9429	22.96	0.9424	23.98
SMB	0.1487	2.17	0.1956	2.86	0.1069	1.63
HML	-0.037	-0.57	-0.0104	-0.16	0.0525	0.85
RMW	0.2136	1.84	0.1102	0.95	0.0955	0.86
CMA	0.1442	1.45	0.2051	2.06	0.2144	2.25
Observations	192		192		192	
R-squared	0.8		0.81		0.82	
Annualized Alpha	4.1%		3.95%		3.38%	
Difference in Alpha			-0.15%		-0.72%	

Table 10.3 Value-weighted decile difference between material, immaterial and Thomson Reuters ESG score

	Material		Thomson Reuters		Immaterial	
	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0009	0.68	0.0016	1.04	0.0017	1.43
Mkt_RF	0.929	25.31	0.9515	22.58	0.8814	25.99
SMB	-0.0531	-0.88	-0.0771	-1.11	-0.0995	-1.78
HML	0.0009	0.02	-0.007	-0.11	-0.0278	-0.52
RMW	0.1104	1.1	0.0646	0.56	0.139	1.5
CMA	0.1004	1.21	0.1813	1.91	0.2641	3.45
Observations	204		204		204	
R-squared	0.82		0.78		0.81	
Annualized Alpha	1.08%		3.19%		2.31%	
Difference in Alpha			0.81%		1.02%	

Table 10.4 Value-weighted quintile difference between material, immaterial and Thomson Reuters ESG score

	Material		Thomson Reuters		Immaterial	
	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0013	0.96	0.0016	1.41	0.0005	0.42
Mkt_RF	0.9864	26.23	0.9376	29.6	0.9245	26.66
SMB	-0.1024	-1.65	-0.0612	-1.17	-0.0982	-1.72
HML	0.0534	0.9	0.0142	0.28	-0.0041	-0.07
RMW	0.0514	0.5	0.1354	1.57	0.2327	2.46
CMA	0.1378	1.62	0.1931	2.7	0.3056	3.91
Observations	204		204		204	
R-squared	0.83		0.86		0.82	
Annualized Alpha	1.55%		3.95%		3.38%	
Difference in Alpha			0.38%		-0.93%	

Table 11.1 Equal-weighted decile performance on double sort portfolios

	HMLI		HMHI		LMHI		LMLI	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0027	1.78	0.0029	1.94	0.0021	1.38	0.0021	1.36
Mkt_RF	0.9118	21.93	0.9200	22.46	0.9327	21.98	0.9328	21.72
SMB	0.1447	2.09	0.1275	1.87	0.1143	1.62	0.1530	2.14
HML	0.0238	0.36	0.0479	0.74	0.0961	1.44	0.0362	0.53
RMW	0.1244	1.06	0.1114	0.96	0.1118	0.93	0.1211	1
CMA	0.2157	2.14	0.2026	2.04	0.1516	1.48	0.1436	1.38
Observations	192		192		192		192	
R-squared	0.8		0.81		0.8		0.8	
Annualized Alpha	3.26%		3.52%		2.58%		2.57%	
Difference in Alpha			0.25%		-0.68%		-0.69%	

Table 11.2 Equal-weighted quintile performance on double sort portfolios

	HMLI		HMHI		LMHI		LMLI	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0024	1.67	0.0029	1.98	0.002	1.31	0.0018	1.28
Mkt_RF	0.9359	23.58	0.9275	22.61	0.9550	23.15	0.9656	24.61
SMB	0.1192	1.8	0.1231	1.8	0.1229	1.79	0.1289	1.97
HML	0.0239	0.38	0.0291	0.45	0.0388	0.6	0.0301	0.49
RMW	0.1471	1.31	0.1352	1.17	0.1487	1.28	0.1616	1.46
CMA	0.1446	1.5	0.1797	1.81	0.164	1.64	0.143	1.5
Observations	192		192		192		192	
R-squared	0.82		0.81		0.81		0.83	
Annualized Alpha	2.92%		3.59%		2.37%		2.21%	
Difference in Alpha			0.66%		-0.55%		-0.71%	

Table 11.3 value-weighted decile performance on double sort portfolios

	HMLI		HMHI		LMHI		LMLI	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0017	1.3	0.0030	2.18	0.0017	1.37	0.0001	0.11
Mkt_RF	0.9216	25.76	0.8916	23.62	0.9603	28.75	0.8797	24.36
SMB	-0.0796	-1.34	-0.0934	-1.49	-0.1246	-2.24	-0.0614	-1.02
HML	0.0771	1.37	-0.031	-0.52	0.0601	1.14	-0.0438	-0.77
RMW	0.0109	0.11	0.0147	0.14	0.1116	1.18	0.1524	1.49
CMA	0.3625	4.18	0.1402	1.53	0.1236	1.53	0.2683	3.07
Observations	192		192		192		192	
R-squared	0.83		0.8		0.86		0.8	
Annualized Alpha	2.05%		3.64%		2.01%		0.17%	
Difference in Alpha			1.6%		-0.04%		-1.88%*	

Table 11.4 value-weighted quintile performance on double sort portfolios

	HMLI		HMHI		LMHI		LMLI	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0009	0.77	0.0021	1.75	0.0008	0.68	0.0008	0.66
Mkt_RF	0.9147	28.66	0.9230	28.39	0.9633	31.26	0.9025	27.48
SMB	-0.1159	-2.18	-0.0554	-1.02	-0.1213	-2.36	-0.0933	-1.71
HML	0.0409	0.81	-0.041	-0.8	-0.0096	-0.2	-0.0345	-0.67
RMW	0.0014	0.02	0.0776	0.84	0.1296	1.49	0.1803	1.94
CMA	0.2456	3.18	0.2567	3.26	0.1412	1.89	0.2865	3.6
Observations	192		192		192		192	
R-squared	0.85		0.85		0.87		0.84	
Annualized Alpha	1.08%		2.51%		0.91%		0.95%	
Difference in Alpha			1.43%		-0.17%		-0.13%	

Table 12.1 Comparison performance firms with and without ESG data

	Value-weighted				Equal-weighted			
	With ESG data		Without ESG		With ESG data		Without ESG	
	coeff	t-stat	coeff	t-stat	coeff	t-stat	coeff	t-stat
Constant	0.0041	2.64	0.0088	3.97	0.0063	3.53	0.0133	6.99
Mkt_RF	0.8886	20.68	1.0097	16.60	0.8835	18.07	0.8362	16.02
SMB	-0.0665	-0.92	0.1330	1.30	0.1614	1.96	0.2827	3.22
HML	0.0145	0.20	-0.1631	-1.59	-0.0478	-0.58	-0.1013	-1.15
RMW	-0.0005	0.00	-0.0697	-0.39	0.1364	0.95	-0.1857	-1.21
CMA	0.1653	1.56	-0.1216	-0.81	0.0969	0.80	-0.0715	-0.56
Observations	144		144		144		144	
R-squared	0.80		0.76		0.78		0.76	
Annualized Alpha	5.09%		11.12%		7.83%		17.23%	
Difference in Alpha			6.03%***				9.39%***	

Appendix 6: Robustness table for immaterial issues

Table 9.1. Robustness Tests Investments in Immaterial Sustainability Issues

	Value-Weighted			Equal-Weighted		
	Low Decile	High Decile		Low Decile	High Decile	
	Annualized Alpha		Difference	Annualized Alpha		Difference
Alternative factor models						
Raw return	9.34%	10.86%	1.52%	9.6%	9.75%	0.15%
Three factor	3.05%	4.03%	0.98%	3.13%	3.07%	-0.07%
Four factor with momentum	3.39%	4.17%	0.78%	3.66%	3.42%	-0.24%
Alternative weights						
Unweighted materiality matrix	0.84%	3.1%	2.27%*	2.49%	3.12%	0.62%
Subset of firms						
Excluding Human capital firms	2.25%	3.88%	1.64%	2.67%	2.87%	0.21%
Excluding polluting firms	1.59%	2.69%	1.1%	0.91%	1.68%	0.78%
Excluding below mean small firms	1.55%	3.16%	1.61%	1.58%	3.07%	1.5%
Subset of periods						
2004 –2011	2.93%	3.09%	0.17%	4.91%	3.94%	-0.97%
2012 –2019	-0.19%	0.78%	0.97%	-2.2%	-0.99%	1.21%