Master Thesis
Copenhagen Business School

2017

Responsibly active

How institutional activism through ESG-engagements affect firm performance and investor returns

Authors

Andreas Åström
MSc Applied Economics and Finance

Robin Gaastra
MSc Applied Economics and Finance

Supervisor: Thomas Einfeldt
Page Count: 120 (136 w/ appendix)
Number of Characters: 272 730

Acknowledgements

We would like to extend our great appreciation to Schroders for making their internal data and information available to us and for participating in interesting discussions throughout the process of writing this thesis. We would also like to thank our supervisor Thomas that has, in a fun and engaging way, helped us stay on course and offered interesting and important insights and comments.

Executive summary

Based on a private data set of active engagements on topics around Environmental, Social and Governance, this thesis identifies that certain ESG-engagements have higher likelihood of being successful both in terms of completion and of value creation but that there generally is a negative to neutral effect of a firm being engaged in terms of improved Excess return, ROA, Tobin's Q and Sales per employee. With the data set provided by Schroder Investment Management, this paper takes an exploratory approach to investigate how active ownership on ESG-issues is received by the target firm and how it affects operational performance measure and stock market reactions in the time period 2005-2016. By first assessing objective engagements, the paper finds that there are great differences between ESG-components, sectors, topics, regions and more and how likely it is that the objective is achieved by the firm. With less than half of the objective engagements being achieved, this finding is of large practical importance as it indicates that certain engagement specifications are more successful and deemed more important than others, and that the differences are unneglectable. By investigating the process one level deeper than most existing research, the finding that there are no significant relations between ESG-engagements and the corporate performance measures ROA, Tobin's Q and Sales per employee, adds value to the discussion about causality between ESG-initiatives and corporate performance. The lack of significance is discussed and it is implied that practitioners need to relate the engagement objectives more closely to observable and measurable areas in order for the efforts to be priced appropriately. Lastly, the paper investigates if the goal of the active owner is met, i.e. to generate excess returns. A negative to neutral excess return is found which indicates that the efforts are not currently translatable to increased shareholder returns. Importantly, the model finds that the market reactions improve when only the achieved objective engagements are analyzed, further strengthening the importance of more specific, detailed and cohesive engagements efforts that are likely to be achieved, measurable and valuable. Taken together, the paper complements existing research in analyzing the value chain of active ownership on ESG-issues but also offers insights for practitioners that can allow for more accurate engagements going forward.

Table of contents

List of figures

List of tables

1. Introduction

The initial chapter will provide the reader with background information on the subject of the thesis and introduce the topic broadly. Having established some familiarity with the topic itself and some current issues within it, the focus will be narrowed to describe the purpose and objectives of this paper. Together with the problem formulation and a discussion around how this paper can contribute to the field of research, the limitations to conducting this study and the strategy and design of the process will be presented. The chapter attempts to provide the reader with an introduction to the subject and an explanation of how it will be approached.

1.1 Introduction

Practically, whichever measure is preferred, the interest and attention for responsible investment and more particularly, ESG-investments has increased significantly. With $ 6,57 trillion in Assets Under Management (AUM) in 2014 in the U.S and a European market that from 2013 to 2015 increased in size by 39%, the traction to and interest in ESG-investments has increased significantly (Deutsche Asset & Wealth Management [Deutsche Bank], 2015). ESG, being an acronym for Environmental, Social and Governance, encompasses investment activities that account for these three factors in investment decisions, to manage risk and long term returns (United Nations Principles for Responsible Investments [UNPRI], n.d). The combination of the three constructs a measure of sustainability and ethical impact related to an investment. Yet, with approximately 1500 studies on the topic, published within the last 15 years, there is still no consensus around the performance of such investment strategies (Deutsche Bank, 2015). However, it does stand clear that responsible investments and the integration of an ESG-criteria in investment processes and strategies is gaining momentum across continents and is becoming more mainstream and easily accessed for both asset managers, private investors and company managers. It has also been realized that activities and operations that are considered in line with good ESG-practices have been starting to affect financial valuations of companies and their projects, in the sense that the financial industry has understood the importance of being able to value such activities (European Sustainable Investment Forum [Eurosif], 2016). This increased focus has been spurred on by investor demands as private and institutional investors have been searching for ways to invest with non-financial aspects in consideration. Parallel to this, the existence of "universal owners" is becoming more prevalent. The impact of such investors is large both in an economic sense but also in non-financial terms and this interrelation has become

increasingly important for these investors as the actions of one of their portfolio companies potentially affects another company in the same portfolio (Mattison, Trevitt & van Ast, 2011). With such considerable amounts of assets, universal owners are greatly affecting what trends and tendencies are explored also by the broad public, something that has worked in favor for ESG-integration. Nonetheless, there are still discrepancies between theories and proof and it can be argued that business and valuation models are not able to accurately value and assign costs for off balance sheet actions such as environmental damage, socially harmful operations and poorly structured governance systems.

This discrepancy has been addressed by a significant body of research and a large number of measures proving positive relations between ESG, corporate financial performance (CFP) and financial returns have evolved (Deutsche Bank, 2015). There is a rather clear opinion across approaches that there is a positive relation between ESG and CFP and that this is the case also for the ESG-components individually. Nonetheless, it took quite some time for regulatory bodies to acknowledge these potential benefits and the fiduciary rights of investors was for a long time a setback for ESG-integration. As such, investments were not deemed financially comparable to "normal" investments due to the lack of proof for the financial attractiveness of ESG-integration. With the opinions and attitudes towards ESG-integration starting to go from skeptical to more encouraging, investors naturally also started to channel funds to organizations addressing these areas and numerous criteria and ratings were developed to keep track of the ESG-quality of a firm. One of the most commonly used measures today is the MSCI ESG rating where firms can be compared to their industry peers and, much like bonds, receive a rating on a score AAA-CCC (Morgan Stanley Capital International [MSCI], n.d). What this meant was that investors could efficiently introduce screens into their portfolios and ensure a certain ESG-quality across the investments. Ultimately, this reduces the investment universe and harms diversification and thus the overall performance of the portfolio. Being an unwanted feature, there has to be an excess reward in doing this, which, as mentioned above, has been proven to be the case in a lot of situations. But what becomes interesting is how investors, through active interactions with the portfolio firms can accentuate ESG-initiatives that are seen as value enhancing. It is activities like this that large institutional investors, like Schroders, have initiated in order to inspire, encourage and assist their portfolio companies to undertake ESG-activities with the aim to see higher returns on their investments at the end.

It has been proved to exist a relation between ESG-activities and numerous performance measures and financial benefits for firms. Also negative and neutral results have been highlighted in some

relationships and an overall consensus is not completely formed. This thesis will take the analysis and measures one level deeper by looking at how *active engagements* from the institutional investor, Schroders in this case, can influence value creation through ESG-activities and how this relates to superior returns, improved operational performance and similar favorable characteristics. In doing this the benefits are plural. Firstly, Schroders receive deeper insights into how successful and efficient their efforts have been in relation to their goals. Secondly, other investors can more thoroughly evaluate ways that firms can undertake ESG-activities and how these might be best encouraged and included into the operations of the firm. Thirdly, it complements research in painting a broader understanding of how ESG-activities on a firm level can influence investors and stakeholders. Lastly, the results will identify how well aligned the actions of investors are with the subsequent firm performance and stock market valuation on topics within the ESG-spectra.

1.2 Problem Discussion

Defining one or a few specific issues or concerns that captures the wide array of topics within the ESG-umbrella can be detrimental. This project will take a holistic perspective in approaching the hypotheses posed and will drill down into further analysis where it is deemed necessary and insightful. This is seen as important and suitable given the current knowledge on the area. Being able to compare effects between the specific ESG-components and also to the aggregated results will allow for a richer discussion and ultimately greater insights. As a result of combining an evolving field of research with private data, this paper is to be considered as highly exploratory in the sense that it is not looking to confirm or address specific relations but rather to identify areas and relations that can be useful for both practitioners and researchers.

1.2.1 Research Objective

This study has dual objectives, both of which requires a solid background of theoretical understanding. Theories provide a sound point of reference when analyzing and interpreting the results of the study itself and also builds towards a consensus about what types of methodology is appropriate to use in this field of research. To strengthen the practical usefulness and understanding of the issues, the theoretical insights will be complemented by information from practitioners that has been produced for investors, institutions and the broader public. The first objective of the paper lies within this area; to further broaden and deepen the understanding of investors, practitioners, institutions and researchers about the relation between ESG-engagements, the effects on firm performance and ultimately the returns to investors. The second objective is directed at the practitioners at Schroders

and to strengthen their knowledge about the value they create through their engagements with portfolio companies. This will be done in multiple ways, which will be discussed more in detail below. The first objective will build of the theoretical background provided in chapter two and be reiterated in chapter five and to some extent six where the results from this paper will be presented and analyzed. The second objective is dependent on but not defined by the observations and insights presented from the theoretical review. Partly because this study aims to complement previous research and partly because the tests and results addressing the second objective has not previously, to the knowledge of the authors, been conducted with the same type of data and methodology thereby creating unique perspectives. Ultimately, the aim is to present ideas and thoughts of what this means for the operations of Schroders and how this can be leveraged going forward.

1.2.2 Problem Formulation

Although ensuring a holistic perspective is important, defining more specific problems to be addressed is necessary to direct the study and create a structure within which the paper can be elaborated freely. The main problem of this thesis lies within the understanding of how well the responsible activities of firms, in the areas of Environment, Social and Governance, translates into their 10k-report or national equivalent and ultimately the ability to produce excess risk adjusted returns. More specifically, the problem statement of this thesis is;

To measure how effective the active ESG-engagements from Schroders are, how they are affecting the performance of the target firm, how this translates to the stock market and what this implies for the area of ESG-investments as well as for Schroders.

This will be accomplished through three approaches;

1. Identify what types of engagements have been the most successful in terms of being achieved and carried through by the target firm.
2. Estimate how an engagement might lead to a change in operational performance, firm valuation or productivity.
3. Quantify how these engagements affect the stock price of the target firm and how the return to Schroders vary with their engagement efforts.

After establishing these items, the implications and effects that this has on Schroders' operations will be highlighted and brought to light in terms of how it can be further improved and conceptualized.

1.2.3 Limitations

The implications and analytical boundaries of this paper is strictly determined by the reach of the data set. In working with private data there is a huge gain in practical application and usefulness, but a slight loss in the ability to generalize the results to a broader population. Although continuously applied and contrasted to existing theoretical works, the results will be interpreted from a Schroders-perspective which somewhat hampers the practical recommendations it can bring to existing theoretical body. To some extent, the exploratory stage that research on this topic is at still limits the possibilities to refine and perfect quantitative methods and measurement techniques and puts greater importance on exploring potential relations and connections. Although an exploratory approach is highly suitable for this thesis, the methodology still offers creativity and individual techniques to define the work, something that might be considered to further limit the external validity of research. Taking a quantitative approach separates the paper and the results from the qualitative aspects of responsible investments. Such are important and to a large extent they define the topic, however the inclusion of interviews, expert opinions and general ideas and thoughts amongst investors would alter the focus and more explicitly become a report of use for Schroders, but where the connection to theory and wider application has been deemed important.

1.2.4 Research Strategy and Design

This paper will have a positivistic epistemological approach, implying that the information considered is broadly available and has been derived objectively without the research being influenced by the researchers own values. Data or facts will only be provided and included having considered its external observability and the process is defined in a way that allows for out of sample generalization as much as possible (Saunders, Lewis & Thornhill, 2016). As elaborated upon earlier, this paper has dual objectives. To best meet these both a descriptive and an exploratory approach will be applied. The descriptive approach to the theoretical review is motivated by the importance of creating a clear picture of ESG-engagements. Descriptive research attempts to *"portray an accurate profile of persons, events or situations"* (Robson, 2013) and in doing this, the approach offers an understanding of the potential of ESG-engagements. The exploratory purpose will seek to find out *"what is happening, seek new insights, to ask questions and to assess phenomena in a new light"* (Robson, 2013). An exploratory study is considered useful when the objective is to clarify the understanding of a problem when the precise nature of a problem is not clear. Since ESG is perceived as a quite complex and abstract field of research, the exploratory study is motivated to provide further insights to the managers at Schroders. This approach is advantageous in a way that it is flexible, the initial focus is typically wide and becomes narrower as

the research progresses (Saunders et al., 2016). Moreover, the approach can also provide us with deeper insights on whether the questions in this project needs to be investigated further. In addition, an exploratory approach is supported by the goals from Schroders as they wish to expand their knowledge of how ESG engagements are actually affecting the target firms and subsequently their own returns.

With an exploratory approach like this, with the benefits expressed above, choosing a methodology that allows for multiple areas of testing, interest and focus is central. This paper will apply a number of different statistical methods in order to quantitatively assess the various areas of interest that comes with the engagements. Having the method and specific test procedure vary between different areas is crucial in achieving the broad and holistic perspective defined in the problem formulation. This distinction leaves the implications to be drawn from this paper more general in the sense that they are not presenting or proposing effects of certain magnitude or results necessitating immediate actions. The attention to detail in the models is kept high, yet the models themselves are not constructed to generate estimates to enable specific and narrow recommendations or results. They are rather constructed to be able to identify relations that are of practical and theoretical importance, as opposed to being of certain statistical magnitude.

2. Theoretical Framework

To better describe the importance of having a theoretical framework the points made by Johnson & Gill (2003) are useful: *"Theory is clearly enmeshed in practice since explanation enables prediction which in turn enables us to assert control over what happens or does not happen - or at least it proffers the potential for doing such things"*. By saying so, it is evident that theories are means by which we generate expectations about the world. Theories are often build upon the perception of what have happened before and thus influence how we set about future interactions with our world. Despite having a more applied nature in this project, a theoretical framework is still seemed to be vital in order to create a structure and systematic process for the analysis. This chapter will provide the reader with some background information regarding this topic, definitions, its evolvement, current situation and the issues it has today.

2.1 Theories of the Firm

Central to any discussion about firm value creation is the debate on what the ultimate purpose of the firm is. Milton Friedman (2002) was amongst the most influential in this field as he, in his book Capitalism and Freedom from 1962, clearly explained why the firm should only care about making profits and distributing this to its shareholders. In this traditional theory *"There is one and only one social responsibility of business - to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game…"*. Opponents of this view tend to focus their critique towards the first part and argues that it is a narrow and perhaps uncivil standpoint. Proponents, and indeed Friedman himself, would instead focus on the last part, the rules of the game. The idea being that if society, governments and regulators truly care about the social impact of firms, then they have all the means to direct corporation's actions in a "responsible" direction through laws and regulations. Thus social responsibility is not the responsibility of each firm, but the government and public.

Two contrasting theories is the "Transaction cost economics" and "Agency theory". The former adopts a contractual approach to describing and analyzing economic organizations and finds transaction costs, and the task of minimizing these as central. First, transaction costs in markets are argued to define the organization and the need for one, as some transactions can be done cheaper, faster and more efficiently within the organization than on the market place. Ultimately this is why there are organizations. Secondly, transaction costs is the main purpose of the firm. More specifically to minimize these; "*...men in general, and within limits, wish to behave economically, to make their activities and*

their organization "efficient" rather than wasteful." - Williamson (1991). An interesting and, to some, appealing feature of this theory is that it allows economic agents to be self-interested, and thus includes opportunism, moral hazard and the agency theory into the picture (Williamson, 1991). The latter builds upon a fact that Adam Smith pointed out in the famous Wealth of Nations (1776) which is that people (managers) will not watch over other people's money (the firm's money and assets) with the same anxious vigilance as they would watch over their own. Based on this idea came what would later be defined as agency costs by Meckling and Jensen. Their theory of the firm recognizes the firm as a black box in a theory of markets where firms are important players. The firm is merely a box operating as to meet the marginal conditions of the outside environment with respect to inputs and outputs, through which it maximizes output and present value. In this sense their approach reminds much of what has been noted earlier by both scholars as Adam Smith and Friedman above, however they add to this a human element that alters the equilibrium of the firm. They account for conflicting objectives of the individuals inside and outside the black firm box and analyses how this changes the potential of organizations and they explain why a failure to maximize the value of the firm is consistent with efficiency (Jensen, Meckling, 1976). These three examples all have varying definitions of what the firm is and how its components interact but to a certain extent agree that the firm is opting to maximize its own profits.

The general public's perception was slightly altered by the emergence of the stakeholder theory of the firm that can be seen as a competing theory to the shareholder view (Pfarrer, 2010). First defined by R. Edward Freeman (2010) in his book Strategic Management: A Stakeholder approach from 1984, the stakeholder theory of the firm, as can be deducted from the name, broadens the purpose of the firm from only maximizing the wealth to the shareholders, to numerous stakeholders such as shareholders, creditors, employees, customers, suppliers, communities and society at large. Tirole (2001) pinpoints the arguments of this theory in mentioning that *"Managerial decisions do impact investors, but they also exert externalities on a number of "natural stakeholders" who have an innate relationship with the firm… There is no denying that such externalities may be substantial…"* upon which the author reflects on whether it can be justified to grant all the power to the shareholders only. This concept is taking a broader view of the purpose of the firm, one in which non-investing parties would be better represented. Jumping in time to more current practices, this idea has been widely adopted as recognized in Accenture's UN global Compact CEO study (2016) where their interaction with more than 1000 CEOs worldwide entails that 89% of the CEOs think that commitment to sustainability is translating into real impact in their industry and 80% believes that demonstrating a commitment to societal purpose is a

differentiator in their industry. Underscoring the ideas that a stakeholder perspective is valuable to the firm, the report affirms that this is a viable and profitable purpose for a firm. This idea does to a great extent match the profile of this report; a stakeholder view of the firm clearly aligns with the practices and investment philosophies that are to be analyzed and is the most applicable for this purpose.

2.2 Investor Trends

Responsible investment attitudes have been present on markets for a long period of time, but seem to have constantly evolved. Deutsche Bank's Climate Change Advisors (2012) presented a report that reviews the development of sustainable investing in a broad setting. They construct three time periods with different characteristics; between 1960s - mid 1990s was the era of early Socially Responsible Investing (SRI). This period was characterized by values-driven investments, ethically-oriented features that primarily took account of corporate social, ethical and environmental actions. The subsequent period, from mid 1990s - present, shifted to a more agile approach where social, environmental and corporate governance issues started to be considered in both the investment decision-making and strategic process. An important shift and distinction to make is that *"current SRI employs a mix of negative (values-driven) and positive (risk and return driven) screening techniques to maximize financial return within a socially aligned investment strategy."* The inclusion of positive screening together with an increased level of shareholder activism has been detrimental in the improved ability for investors to avoid compromising diversification in their search for specific investments. Dated to 2003, the "ESG-era" ensued with the UNEP Finance Initiative (UNEP FI) announcing the commissioning of extensive research of the link between ESG issues and security valuation. This sparked the demand for more rigorous and carefully defined frameworks and guidelines for such engagements and in particular the historical risk-return payoff of SRI was under investigation.

Much of this research was able to identify ESG-investments being in a growing phase and after numerous studies and plenty of years, something of a consensus about the financial contributions of ESG activities started to form. Del Guercio and Tran (2012) dedicate a chapter of their book to introduce and discuss the similarities between the issues ESG proponents are facing today with institutional investors and pension funds, in relation to corporate governance issues about 25 years ago. Except for highlighting the long transition periods of these trends, they note that in 2012 the most important advancement is that institutional fund managers have no legal hinders to incorporate ESG principles into their investments, to be seen in relation to pension fund managers who at that point still were strictly bound under their fiduciary duty. By realizing this, and by arguing that some

investors accept a possibly lower financial return in exchange for positive influence on social and environmental changes and attitudes, they conclude that that SRI and ESG-engagements from institutions and fund managers is highly client driven. This idea is supported by the Social Investment Forum Foundation (SIFF) where 85% of the fund managers asked listed "Client demand" as the reason for incorporating ESG-factors into their strategies (SIFF, 2010). Del Guercio and Tran conclude that 25 years of institutional investor advocacy managed to turn corporate governance notions from radical to mainstream and argue that if practitioner interest is a leading indicator, the relationship between ESG-engagements and financial performance is likely to be carefully studied in future research.

Further notice of the growing, yet not mature market for ESG-concerns in the early 2010's can be found in Kotsantonis, Pinney and Serafeim (2016) approach of myths regarding ESG and investment management. They identify four major myths about ESG-investments and contrast these with their perception of reality. Initially one misconception was that firm efforts to address environmental and social issues always bare a cost and that this is equivalent to a reduction in shareholder value. They address this by pointing at the importance of distinguishing between material and immaterial investments. Material, business-related ESG-exposures, those with large potential effects on the long-run value of the firm have been shown to improve both operational performance and excess returns. Immaterial, relatively unimportant but socially popular ESG-issues however tend to record average or below performances. This perception is troublesome, as institutions tend to invest only in firms that already have a "good" ESG-profile meaning that they have already reached a certain level of materiality and thus fails to produce further value-adding improvements (i.e. material). Secondly, by analyzing the total assets under management for ESG-investments, ESG has yet a long way to go to become mainstream. This suggests that the growing demand for ESG-engagements as suggested by SIF had not gained heavy momentum at this point. Thirdly, they bring up the myth that companies cannot choose their investors and all investors have similar (short term) investment horizons, meaning there is no room for management to pursue more sustainability targets. They argue that this is not the reality as much research has proven that different management attracts different investors. Further they argue that, although ESG-data is nowhere close to be as standardized as financial data, there has been tremendous progress during the last few years in increasing the availability and quality of the data. Instead they argue that the challenge is for investors to identify those ESG-factors that are material to financial performance. In an attempt to reduce this gap, this report will shed light on such areas.

The last myth targets the fiduciaries and their reasoning that under their duty of loyalty to protect the financial interest of their beneficiaries, they must consider only traditional economic factors in their valuation models and therefore must exclude ESG-factors commonly viewed as "non-economic". With the substantial amounts of assets that pension plans have under management globally, their interest and strategies tend to lead the way of common investments (Del Guercio & Tran, 2012). Smith (1996) recognizes in an early study that California Public Employees Retirement System (CalPERS) was, and arguably still is, the most widely regarded and accepted shareholder activist on the US equity market. The inclusion in his study is motivated by the openness of the fund's practices, their long history of shareholder activism and the fund's size and the impact that comes with it. Currently the fund holds a staggering $ 308,65 billion in market value and in 2011 the CalPERS board approved the adoption of a "Total Fund process for integrating ESG issues as a strategic priority across CalPERS' portfolio."(CalPERS, 2016). This addresses the importance identified above where ESG-investments are becoming accepted within the frames of the fiduciary's duties and responsibilities.

Almost simultaneously Mattison, Trevitt and van Ast released a report for the UN-backed PRI and United Nations Environment Programme (UNEP) initiative where part of the scope of the study is to identify how environmental damage poses financial risks to economies, companies and investment funds. Research like this attempt to further investigate and identify the positive financial influences that responsible initiatives can have. Their arguments align much with the ones of Heal (2005) in that they argue that current economic and business models fail to recognize the value of ecosystem services and the cost of environmental harm. *"Where the costs of environmental damage, such as pollution, are excluded from the transaction between a buyer and seller, they are largely "external" to a company causing damage and are borne by third parties."* Their conclusion is that if these "externalities" are not reflected in market prices, they will remain externalities and the destructive production patterns will not stop. Encouraging the realization of environmental costs, they are indirectly putting importance on also valuing the benefits in a financially viable way, creating "costs and revenues" of environmental productions and actions. They further agree with Heal that *"The discrepancy between environmental costs on company balance sheets and those paid by others in the economy represents a market failure."* Supporting their arguments, they estimate the environmental costs of a hypothetical large and well-diversified fund. Valued at $20 billion with 50% of the assets in equities, in the Morgan Stanley Capital International (MSCI) ACWI using the same weightings, they reveal that for every $10 billion invested in this index, about $560 million in

environmental costs stemming from the companies in the portfolio is incurred. This opens up the question, who is responsible? Who should bear these costs?

To solidify their message and to increase the influence on institutions, they turn to address the implications that the current situation has on fund managers and their portfolios. When holding a large, diversified portfolio, it is the aim to have equity in different industries, regions and companies. An issue with this is that some environmental costs that are externalized by any company within this portfolio could come to be incurred by other companies within the same portfolio. *"Rising externalities over time at a portfolio level are generally larger than short-term gains from companies that profit from externalizing environmental costs. Accumulating externalities could lower fund returns overall."* A highly illustrative example within the oil industry; Larger profits from oil companies driven by increased energy consumption would potentially drive up the share price and thus benefit investors in the short time horizon. However, as time passes the detrimental effect on the environment through climate change, with rising ocean levels, more intense storms and changes in precipitation would very likely harm other firms within this hypothetical large and diversified portfolio. More expensive water utilities, real estate assets that need more consequent capital investments and rising production costs brought by scarcity of resources are all examples of costs that could work its way down the portfolio and ultimately create lower total returns for the entire portfolio.

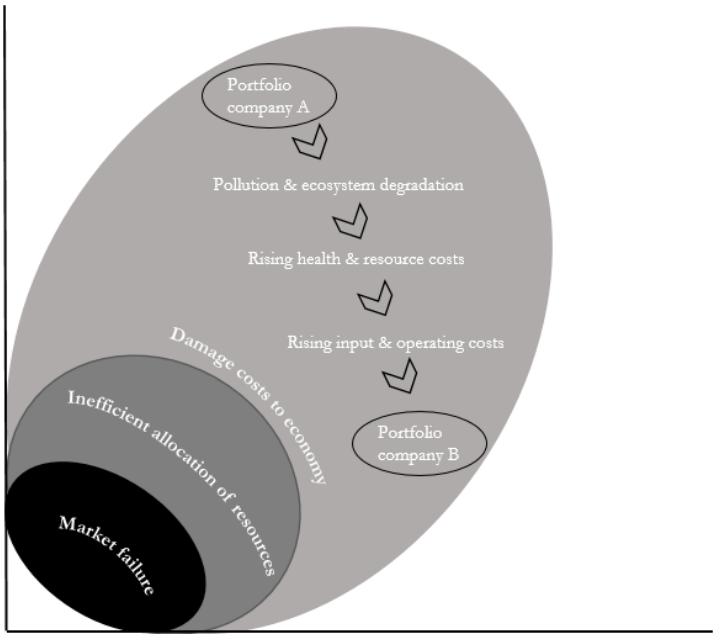


Figure 1: The market failure of environmental externalities. Retrieved from Mattison, Trevitt & van Ast (2011).

As interest and publications have continued to emerge and develop, recent opinions and events project ESG as an almost erupting area. In January 2016, Eurosif together with a coalition of stakeholders, managed to get a joint letter passed in the ECON Committee. This indicated that the Members of the European Parliament (MEPs) *"have acknowledged the materiality of ESG risks for the long-term financial performance of pension funds."* These advancements strive, and pave way, for the general acceptance and

adoption of ESG principals across investor characteristics and purposes. A critical implication from this acceptance stems from acknowledging the evidence of a correlation between non-financial indicators and how it aligns with fulfilling the fiduciary duty of investment managers. Not only does this signal an international recognition of the inherent value that ESG-initiatives can offer, it is also an important attempt to remove the threshold between global investors and ESG-practices. There has also been a fair amount of research done as to how governments, firms and investors can price externalities across value chains. The details of such research will not be thoroughly examined, the interested reader is encouraged to look at for example Tideman and Plassmann (2010), Ding, Zhao, An, Xu & Qian (2015), Bickel, Friedrich, Link, Stewart & Nash (2006) and Soderholm and Sundqvist (2003) for some varying approaches. What these advancements have in common is that they are increasingly directing the public's' and investors' attention to this market imperfection and thus also attracting different solutions.

2.3 Capital Market Theory versus Negative Screening

Closely intertwined with the acceptance and recognition of any investment trend is the potential risk adjusted returns. For responsible investing to fully emerge there are some barriers that needs to be addressed. Barnett and Salomon find themselves, and the research on SRI funds' financial performance, in a deadlock where previous studies are disaffirming one another's results and constantly adding new and alternative interpretations. The opponents of Corporate Social Responsibility (CSR) activities argue that the costs and administrative burden related to those activities have a directly negative impact on the bottom line and thus, firms engaging in CSR activities are in a competitive disadvantage compared to firms that do not. Building on this, rational investors also realize that for a mutual fund (or similar) diversification is of great essence. By applying negative screening and excluding not only potential portfolio companies but also entire industries, fund managers limit their ability to diversify efficiently. Adhering to established portfolio theory concepts, *"the exclusion of firms, industries, and economic sectors has significant implications for the financial performance of an investment portfolio, regardless of its social orientation."* (Barnett & Salomon, 2006). Because of this, fully diversifying the specific risk in a screened portfolio can be argued to be impossible, implying that SRI funds and similar will consistently return a financial loss of various magnitude due to the imperfect diversification. Contrastingly, proponents argue that although the fact that a screened portfolio has a limited pool of companies it could potentially invest in, the pool of firms they do choose from is more capable than the market in general, enabling the companies and the overall portfolio to generate excess returns over time. The support for this argument lies within the stakeholder theory of the firm where

favorable work towards environmental and social factors can help the firm attract and retain employees, lower risk of lawsuits, get favorable attention and support from local communities and is in a better position to get tax breaks from local governments. This builds an immaterial asset that the firm can leverage as a competitive advantage against firms with none or limited engagement in these areas. All taken together, the authors cite previous studies and present the conclusion that *"A simple compilation of the findings suggests there is a positive association, and certainly very little evidence of a negative association, between a company's social performance and its financial performance."* but duly recognizes that *"such a conclusion is illusory. A compilation of findings cannot produce a definitive conclusion given the limitations of the underlying studies."*. To address this, the authors decide to abandon the idea of comparing SRI funds to non-SRI funds with the hypotheses that financial performance does also differ across different levels of screening intensity. Based on and OLS approach they find support for both the portfolio and stakeholder theory. Social screens do narrow the investment choices but if done diligently it can also lead to an increase in financial returns. The funds in their sample that used numerous social screens, efficiently, were on average able to exclude underperforming firms from their portfolios whereas funds with few levels of screening could instead reap the benefits of a well-diversified portfolio. The funds in between the two extremes however, were found to perform worse as they were not able to stock pick the superior firms, nor enjoy the full benefits of diversification.

2.3.1 ESG, Shareholder value and Firm Performance

Despite an increasing prevalence of active ownership, data limitations have left unanswered even what could be considered as basic questions about ESG-activism; how does the market react to engagements and how does ESG-activities affect firm performance? These are some of the questions that Dimson and Karakas & Li (2015) address in their paper. By measuring the cumulative size-adjusted abnormal return the year following an initial engagement, they observe a positive return of 2.3%. If the engagements in addition also have been successful, the cumulative abnormal returns are being positively rewarded with a total 7.4%. The abnormal returns hold true for all the subsamples in magnitude and pattern. By taking a difference-in-difference approach, they are examining the subsequent changes in the targeted firms' operating performance, profitability, efficiency, institutional ownership, stock volatility and governance after successful engagements compared to the unsuccessful engagements. Once again, if the engagements turned out to be successful, they observed improvements in all these measures, meaning that it was an increase in firm performance, investor base and governance whilst the was a decrease in stock return volatility.

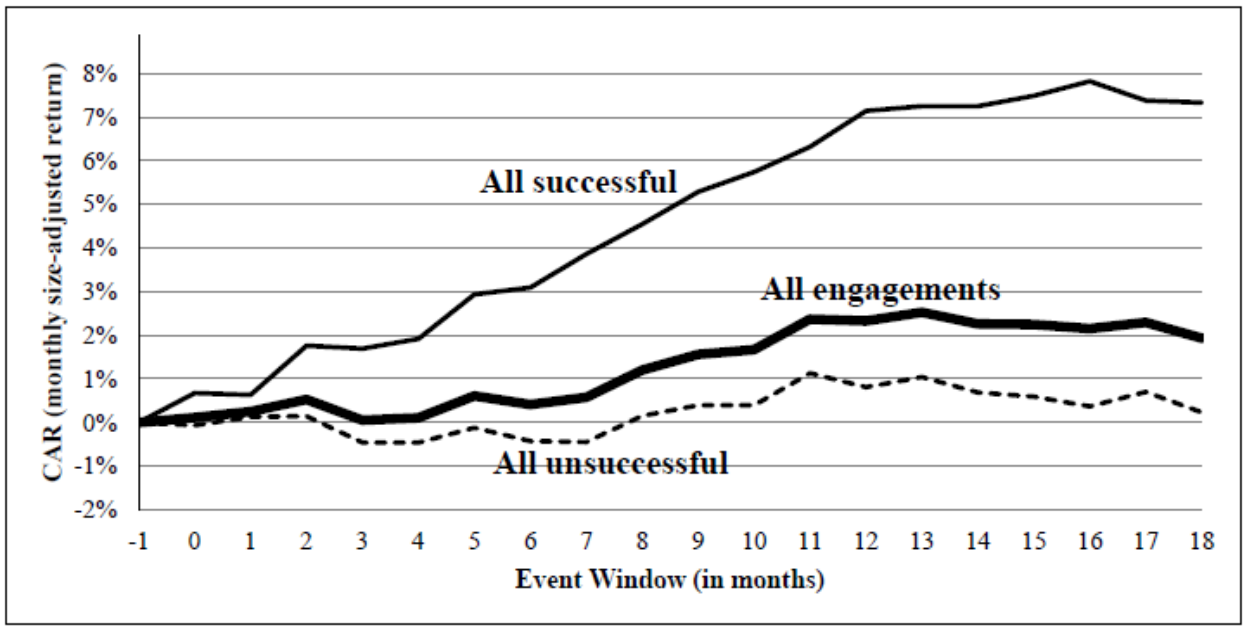


Figure 2: Cumulative Abnormal Returns from ESG-engagements. Retrieved from Dimson et al. (2015).

In an extensive study from December of 2015, a collaboration between Deutsche Bank and researchers at the University of Hamburg (Prof. Dr. Alexander Bassen & Prof. Dr. Timo Busch) was investigating whether integrating ESG into the investment process has had a positive effect on CFP, how a link between ESG and CFP differs across asset classes, if any specific sub-category of either Environmental, Social or Governance had a dominant influence on CFP, and finally, if the effects of integrating ESG into the investment process were stable over time. Their meta-study examines the entire universe of 2,250 ESG-CFP academic review studies that has been published since 1970, out of that, 70% has been published within the last 15 years. Their results reveal that overall less than 10% of the studies displays a negative ESG-CFP relationship, with an overwhelming share of positive results.

The study looks further into to the subcategories E, S & G to find out which (if any) is dominating, by using vote-count studies. Vote-count studies count the number of primary studies with significant positive, negative and non-significant results and "votes" the category with highest share as winner. These type of studies provides robust insights but from a statistical point of view, they are less sophisticated. The results from their vote-count sample of studies show a positive effect similar to the meta-study, but the highest performer of the three is Governance with a score of 62.3% of all studies delivering a positive relation to CFP. Worth mentioning is that Governance related aspects also generated the highest percentage of negative relations at 9.2% from the sample. When they analyzed

various combinations of ESG, only 35.3% reported positive relations, indicating that non-focused approaches seem to lead to a less compelling argument to deploy ESG. Their final investigation was if the ESG-CFP relationship was stable over time. There are theories implying that, due to a growing number of PRI-signatories and the presumption that investment strategies are becoming increasingly more ESG-aware, the correlation between ESG and CFP might diminish over time as a consequence of the apparent existence of learning effects in capital markets. Ultimately this decreases any alpha that could be captured, reducing the attractiveness of such investment strategies. However, from their sample of studies with disclosed correlation factors, there were no sign of a learning curve and they conclude that correlations were stable over time.

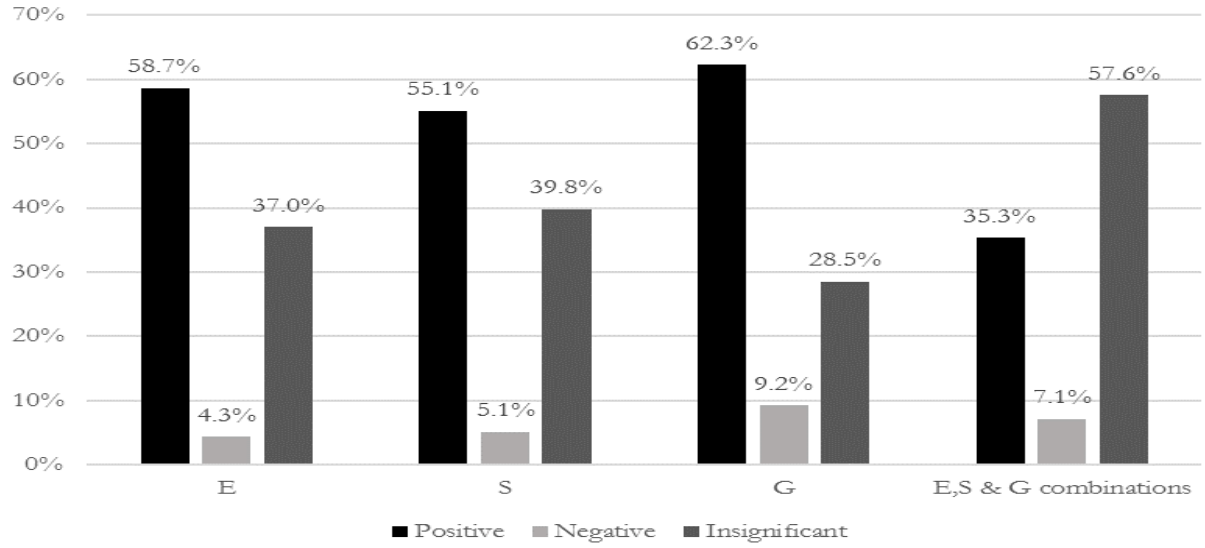


Figure 3: Aggregated results from ESG research and relation to CFP. Retrieved from Deutsche Bank (2015).

The number of studies investigating the relationship between ESG as a whole and CFP is growing. Yet, there are great amounts of research done on more specific topics within each of the three sub-components and as such offers great insights on a more detailed level, a review of the literature existing each ESG-component individually will be conducted.

2.3.2 Environment, Shareholder value and Firm Performance

The relationship between Corporate Environmental Performance (CEP) and CFP is an area of research that has been burgeoning for years and is still expanding due to the attention of the question whether it "pays to be green". Now, given a vast body of CEP-CFP studies, cumulative results have begun to form, catching the attention of recent scholars. By doing a meta-analysis of the yielded findings of event studies assessing the stock market reactions to CEP-related events, Jan Endrikat's

(2016) paper sets out to close the gap in research regarding the stock market impact of positive and negative CEP-related events by synthesizing the previous empirical results. The background for his research is based on the fact that neoclassical economics propose that any discretionary effort toward improvements of CEP will decrease potential profits and thus violate the shareholders' value maximization (e.g., Friedman 1970; Jensen 2002). While other scholars have challenged that view by putting forward arguments saying that it does exist a positive impact of CEP on CFP, the empirical examinations have long yielded mixed results. The findings range from negative relationships, to non-significant effects to findings showing a significant positive relationship between the two variables. Meta-analysis, is a tool that Endrikat uses in order to quantitatively synthesize the findings across a number of studies, and this approach has become increasingly common in management research, especially in cases where inconsistent findings impedes any generalizations for establishing consensus. To answer the question regarding the relationship between CEP and CFP, Endrikat has put together two hypotheses. In the first one, he does an examination across the body of event studies to find out if there is a significant stock market reaction to both positive and negative CEP-events. The outcome of his meta-analysis reveals that there is a positive stock market reaction to positive CEP-related events and negative stock market reactions to negative events respectively. In view of these findings, the second of Endrikat's hypotheses is proposed as following: Across the body of event studies, the stock market reaction to negative CEP-related events is stronger than the stock market reaction to positive CEP-related events. In several event studies, greater stock market reactions (in terms of magnitude of the Annual Returns (AR) or Cumulative Annual Returns (CAR)) could be observed around negative events compared to positive events. Thus, in his paper there is evidence pointing out that the market punishes firms in cases of negative CEP-related events to greater extent than it values cases of positive CEP-related events. In other words, the stock market reactions to CEP-related events appear to be asymmetric. Further support for this type of behavior has been put forward by Klassen and McLaughlin (1996). They point to the fact that negative events such as environmental crises entail financial implications more obviously than positive CEP-related events. Such events are not just signaling poor CEP, it forwards information with manifested financial implications as opposed to positive events, such as announcements of environmental awards, where the financial implications are of significantly lower magnitude. Considering the values gained from analyzing the first hypothesis with support from hypothesis number two, there are multiple effects of CEP-events that can affect firm performance and investor returns.

2.3.3 Social, Shareholder value and Firm Performance

Deutsche Bank (2015) report that out of research looking at the link between social engagements, or CSR, up until 2015, 55.1% finds a positive relation to CFP and only 5.1% report a negative connection. Citing the US SIF report from 2014 they also conclude that the social segment is the largest when looking at AUM from money managers funds, with 34.2% of the capital allocation. The positive relation between social engagements and firm performance is further acknowledged in the article "From Shareholder to Stakeholder" as it reviews current research. They put forward research proving both positive and negative relations, however there is a clear trend that social engagements have a positive impact on both operational performance and stock prices. A few years prior to this Margolis, Elfenbein and Walsh (2007) interpreted the relation between Corporate Social Performance (CSP) and CFP in a meta-analysis. The overall effect they find is positive but small, however the results show that there is no financial penalty for CSP within their sample and they note that there seem to be a link from prior CFP to subsequent CSP that is as strong as the reverse relation. In contrast to the reports above, they find that 58% of the studies show non-significant relations, 27% a positive relationship and 2% a negative relationship between the two variables. The 35 years of studies that have been included in their study has thus been unable to consistently provide results that are economically significant. However, comparing the results of their analysis to the ones drawn above shows an interesting thought; there has been huge changes and improvements in the value creation of CSR-activities. In their analysis, the authors present four broad sets of implications that they believe is characterizing their results:

1. There is a lack of financial reward (penalty) for firms who engage (do not engage) in socially responsible activities. Either the markets do not have an efficient system in place that identifies good and bad behavior and thus enables rewards/penalties or they are not seen as valuable/value destroying enough for the broad investor public. This lack of tangible connection suggests that CFP is not a plausible justification of doing CSP in their opinion and according to their findings.
2. Across their sample, and thus across much of the research conducted, there is a great variation between what sort or type of CSP activity is undertaken. For example, they find a stronger connection for charitable contributions, revealed misdeeds, self-reported social performance and observer perceptions and weaker connections to CFP for other types of engagements.
3. The direction of causality might be reversed, insinuating that there might be a stronger link between how CFP predicts subsequent CSP. In this light they argue that much research

overlooks this aspect and that ultimately understanding how CFP gives rise to CSP should be of more focus.

4. Beside the relationship to CFP, CSP might also be assessed on different parameters; its legitimacy, its value and its effectiveness. That studying and engaging in CSP is legitimate is at this point generally accepted. So far, the value created from CSP-activities can be debated but it is evident that some activities have a stronger contribution to value creation. The efficiency of CSP is yet rather undiscovered and no common tools to measure this have been produced, leaving the field quite open.

Heal (2005) approaches the first point above and evaluates CSR-activities from a reflection on the financial markets and his analysis suggests that there is a resource-allocation role for CSR-programs and argues that where there is a market failure in an industry, that creates a private-social cost differential which is where considerations such as CSR becomes of interest. In a comparison between tech companies and tobacco & oil companies he argues that surely all firms do good; they create jobs, improves wealth and returns for society and thus has a positive presence. However, tobacco & oil firms simultaneously produce a negative presence through health and environmental issues. So where is the difference? Where do we draw the line? *"To understand this we have to see when the interests of corporations are fully aligned with those of society as a whole and when they are in conflict, and for this we have to go beyond Adam Smith, to the concepts of private and social costs."* In this setting, where inconsistencies between firm and social values exist (arguably a market failure), CSR is a detrimental part of corporate strategy as it can improve a firm's operations, staff morale, negotiating power with governments etc. The paper concludes in defining the purpose of CSR as *"...CSR is to anticipate and minimize conflicts between corporations and society and its representatives, aligning private and social costs if differences are the source of the conflict, or minimizing distributional conflicts if these are the issue."* Such a conclusion offers similar interpretations as seen with the relation between CEP and CFP, that CSR-activities might be asymmetric in the relation to CFP.

Brammer and Millington (2008) recognize the importance of measuring specific sorts or types of CSP and their relation to CFP and decides to focus on charitable giving and in doing so they address the second point of Margolis, Elfenbein & Walsh above. Their findings suggest that high and low levels of firm spending on CSP are related with higher financial performance. Further, firms with average social performances mostly enjoy the benefits in the short term whereas the upper segment of social performers could reap benefits for longer periods. Flammer instead studies the effect of CSR

shareholder proposals on financial performance with focus specifically on proposals that are close calls. Such proposals either pass or fail the shareholder vote by a small margin and *"Intuitively, there is no reason to expect any systematic difference between a company for which a CSR proposal passes with 50,1% of the votes and a company for which a similar proposal fails with 49,9% of the votes."* Her primary finding is that the passage of close call CSR proposals significantly increases shareholder value and that abnormal returns seem to converge to zero as the vote moves towards the two majority ends. This indicates that clear-call proposals have already been incorporated in the share price and that only uncertain projects are value generating. With this realization, the author moves on to examining whether there are different ways that CSR proposals increase shareholder value. There proves to be a positive impact on operating performance, and more specifically an increase in labor productivity and sales growth.

McWilliams and Siegel (2000) take the stance towards the third point above as they realize that most econometric models that have previously assessed CSR-spending in relation to firm performance have not controlled for R&D-spending. By recognizing that R&D-spending is an important aspect of firm performance, they include it in their model, to more accurately isolate the effects of CSR individually. They find that most research seems to be upwardly biased in estimating the relation and their "properly specified" model finds that CSR has a neutral relationship to firm performance. The "small" adjustment to the standard econometric of adding variables of R&D and Advertising spending generated different results and indicates that a large body of previous models tested, that has ignored the influence of these two variables, are incorrectly specified and their results questionable. Their argument that most studies present upwardly biased estimates and that when R&D spending is included, the significant relationship vanishes, leads into the argument that the causality might go the other way and that CFP to a greater extent results in CSR than the reverse.

The fourth implication above is addressed by Godfrey, Merrill and Hansen (2009) who test whether CSR-activities can be seen as a goodwill act that has an "insurance-like" property. They hypothesize that CSR-activities creates positive attributions with stakeholders, who in cases of conflict or disagreements with the firm temper their negative judgments and sanctions toward firms because of this goodwill. Expressed differently, they want to assess whether CSR-activities can provide an insurance mechanism to *preserve* rather than *generate* CFP. In their results, they present that institutional CSR-activities, which is those aimed at a firm's secondary stakeholders or society at large, provides this insurance benefit. However, activities targeting a firm's trading partners do not. *"...CSR-based moral capital creates value if it helps stakeholders attribute the negative event to managerial maladroitness rather than*

malevolence, and temper their reactions accordingly." Supporting their results, they refer to the work of Smith and Stulz (1985) and Stultz (2002). Capital market theorists would argue that the investor can diversify every idiosyncratic aspect of the firm and therefore capital spent on risk management cannot add value to the shareholders. Together the three studies show that such violations of the perfect market assumptions do add value as they protect investors against deadweight costs in financial distress in a way that is not possible to achieve through the use of market mechanisms. Hence, these papers are addressing a slightly different and broader value and effect of CSR-engagements. This is built further by Dhaliwal, Li, Tsang & Yang (2014) who examine the benefits associated with CSR-disclosure in an international setting, covering 31 countries. Their approach to assessing CSR-effects is focused on the differences in value creation and efficiency across geographic regions. The report identifies the numerous difference between countries in their legal systems, norms and approaches to CSR-reports, something that confirms the multiple faces and hats of CSR-engagements.

2.3.4 Governance, Shareholder value and Firm Performance

Throughout previous research, governance-related aspects have experienced the greatest positive relation to CFP, but also the most negative occurrences. 62.3% of the covered reports prove a positive relation between the variables whereas 9.2% are negative (Deutsche Bank, 2015). Indicating a greater variation, it becomes interesting and important to attempt to isolate the positive activities and separate them from the negative ones. Gillian and Starks (2000) does just this in discussing the effect of corporate governance proposals on firm variables and specifies different classifications of proposals and measure their subsequent success. One finding, that supports a negative reaction to governance engagements, is that when an institution or a coordinated group sponsors a shareholder proposal, on average, investors perceive that to be negative information. This is believed to be because the appearance of the proposal is signaling management's unwillingness to get involved in the issue, ultimately increasing the risk for conflict. Another broad finding, that institutional or coordinated sponsorship of proposals has a significantly positive influence on the voting outcome, identifies the importance of who is encouraging the initiatives. Further they identify a big difference in the shareholder response to issues relating to antitakeover devices (achieving the greatest support) and proposals targeting executive compensation, director ownership and the limitation of director terms (reaching the lowest support). In general, their report highlights various circumstances that greatly shift the perception and reaction to governance suggestions.

Bebchuk, Cohen and Ferrell (2009) complements this by investigating 24 provisions followed by the Investor Responsibility Research Center (IRRC) and tries to find which of these provisions are playing a key role in the link between corporate governance and firm value. They note that in the timespan 2003-2004 there was four out of these 24 IRRC provisions that stood out in terms of number of precatory resolutions that was submitted and passed;

- Against supermajority provisions, supermajority merger requirements, limits on charter amendments, and limits on bylaw amendments (100% passed)
- Resolutions against classified boards (91% passed)
- Against poison pills (72% passed)
- Against Golden parachutes (62% passed)

Based on these they create an index called the E-index, or only E, which stands for entrenchment level, making a high E-value rather unattractive from a governance perspective. Testing against Tobin's Q and stock returns they find a significantly negative relation between E and the two dependent variables. *"We conclude that, although low-Q firms tended to have high E levels at the end of our sample period, the negative correlation between Tobin's Q and E at the end of our sample period was not all due to the correlation in the beginning of the period; while high E firms began the period already with a lower Q, their Q further declined over time."* To further solidify their results, they test it to a Fama-French and Carhart four factors model where they take long positions in the firms with low E-index scores and short positions in the firms with a high E-index score. Such portfolio strategy, depending on using an equally weighted or value-weighted strategy, realized abnormal returns between 7.4% and 14.8% annually. Based on this, they conclude that the provisions in the E-index should be targeted by researchers as well as by private and public decision makers.

Larcker, Richardson and Tuna (2007) conducts a report in much the same manner where they, based on 39 individual governance measures, construct 14 multi-indicator indices as an attempt to find more reliable and valid ways to measure the complexities of corporate governance. Based on these indices they test for numerous performance indicators and their findings include evidence that for example firms that have a greater proportion of block holders, a compensation mix that is weighted toward accounting performance, lead directors, smaller boards, and fewer busy directors exhibit superior future operating performance as measured by industry-adjusted ROA. To increase future excess stock returns, their study suggests that a compensation mix weighted toward accounting performance, a lead

director and low insider power are important tools. Relating to Bebchuk et al, both studies present similar implications based on two different sets of data and additionally, they recognize the econometric issues and concerns regarding validity in their methods and thus encourages future research to conduct similar tests with various methods to validate the results and to reach a more general consensus.

The relation between governance structures and stock returns and accounting measures is further studied by Carleton, Nelson and Weisbach (1998). They access a private database consisting of the correspondence between Teachers Insurance Annuity Association - College Retirement Equity Fund (TIAA-CREF) and 45 firms it contacted about governance issues between 1992 and 1996 and thus angles their study not only to private information but also towards institutional activism. They recognize the variation in previous studies between institutional activism and stock returns and accounting measures. Their initial finding is that when insider holdings are high, a proposal is less likely to be agreed upon. This is supported by the argument that with higher insider holdings, the firm will be less keen on minding their brand and reputation to outside shareholders. Further, on a more speculative basis, they find a significantly negative CAR surrounding the targeting dates (the date when TIAA-CREF initiated contact with the target firm) in matters of board diversity targets. This relation switches and they see a significant positive CAR for blank check preferred targets and then CAR is insignificantly affected surrounding confidential voting targets. Perhaps the most important aspect pointed out, with regard to the experienced inconsistency in results is *"This emphasizes that previous studies based on public information understate the magnitude of institutions' attempts to influence management because they potentially exclude a substantial number of cases in which management adopts the suggested changes without public knowledge."* Black (1998) adds on to this in his survey among institutional investors in the US and relates their corporate governance activism to subsequent firm performance. The premise of his work builds on earlier studies and he recognizes that institutional activism should not be categorized as a bad thing but it surely cannot substitute for vigorous corporate control market.

Corporate governance has been argued to have influence over several other types of firm characteristics and aspects. Edmans, Zur and Fang (2013) study the effect of stock liquidity on block holders' choice of governance mechanisms. Their findings suggest that liquidity does not dissuade block holders from governing altogether, but instead encourages them to govern through exit rather than voice. Earlier literature has also specifically investigated the effects of corporate governance on cost of debt, and the conclusions are relatively clear: good corporate governance pays off in terms of

reduced borrowing costs (i.e., credit spreads). It has been documented that certain governance measures have a significant impact on a firm's cost of debt, for example the degree of institutional investor ownership and shareholder rights but arguments against this has been proven with comparable strength (Bhojraj & Sengupta, 2003; Cremers & Nair ,2005; Schauten & van Dijk, 2011; Ashbaugh-Skaife, Collins & LaFond, 2006). It has also been shown that corporate governance affects the cost of equity as Ashbaugh-Skaife, Collins & LaFond (2004), show that well governed firms exhibit a cost of equity financing 136 basis points lower than their poorly governed counterparts. Even after adjusting for risk, the difference between well-governed and poorly governed firms is 88 basis points. This is not surprising, as good corporate governance translates into lower risk for corporations and reduces information asymmetries through better disclosure. Taking the cost of debt and equity together, Barth, Konchitchki & Landsman (2013) find results that show how greater corporate transparency with respect to earnings significantly lower a firm's cost of capital.

2.4 Market Failures

As noted above the inability to efficiently, consequently and coherently price externalities and corporate misbehaviors is not only interfering with the efficient market hypothesis but also increases the systematic risk across markets without having investors properly rewarded. It can be argued that the inability to consequently value and price companies' efforts to reduce the gap between corporate profits and societal goals is today's most challenging market failure. In his exhaustive publication "The Stern Review (2006)", Sir Nicholas Stern, address these concerns with a great urgency as he deems climate change as the greatest market failure ever seen and one of the key messages is that *"Climate change is a result of the externality associated with greenhouse-gas emissions - it entails costs that are not paid for by those who create the emissions."*

Mattison et al. (2011) exemplifies how the lack of pricing for externalities comes into play. The issue stems from the inadequate amount of data on how environmental impact is internalized into firm's books and activities and as a consequence it weakens the efficient market hypothesis, where all available information should be reflected in the share price. *"The focus of equity markets on quantifiable, near-term influences on financial performance contributes to lack of transparency in how externalities pass between the economy, private enterprise, capital markets and investors."* To bypass this lack of data, the authors use a theoretical framework that creates a relationship between institutional (universal owners) equity portfolios, externalities, Gross Domestic Product (GDP) and companies' future cash flows. Relying on economic theory and empirical evidence it has been proven that an economy's future cash flows

can be approximated as the economy's future GDP. Universal owners have large holdings across the economy and can therefore be seen as owning a certain share of the economy's cash flows, and consequently the GDP. Lastly as relatively portion of an economy's GDP can be derived from corporate value creation, enabling us to *"see that the long-term price of a universally-owning institutional investor's portfolio represents the universal owner's part of the appropriately discounted sum of all future GDP proportions of corporations…"* This concludes that externalities are likely to have a significant impact on an institutional investor's portfolio future cash flows, with a potential effect of 11-18% of the amount. Levitating this further, where externalities are priced and incurred as costs by the responsible organization, this reduces investments in growth and operations, which in turn might lower dividends and asset values over time. *"We see the universal ownership theory as an absolutely essential part of our investment philosophy - addressing externalities is crucial. Markets that are not working properly destroy value for participants and have inefficiencies. If a company is constantly externalizing costs it is less efficient than its rivals."* - Paul Lee, Director, Hermes Equity Ownership Services. In doing this, they urge investors, researchers and scientists alike to target the companies, industries and activities that represent the worst externalities with the aspiration of creating ripple effects encouraging or forcing improvement upon the entire economy. It is with this in mind, and the identity of Schroders as one of the largest asset managers in the world, that research and papers like this one is of great value. Surely for the individual firm but preferably also for a greater audience of investors, influencers and people able to accomplish change in the area.

3. Methodology

The purpose of this chapter is to walk through the general approach that has been taken in relation to the data set provided but also to make the reader familiar with the definitions and structure of the data set. There will be a section addressing the data from Schroders where the focus will be to present the opportunities but also the limitations in what can be conducted in terms of testing and methodology. To follow, the additional data that has been used in the paper, coupled with the data from Schroders, will be presented briefly. To, on an even deeper level, familiarize the reader with the data, some descriptive statistics has been included to give a fuller picture of the variables and their characteristics and the engagements themselves. Once a solid understanding of the data has been achieved, the second half of the chapter will relate this to aspects regarding reliability and validity and how the data used in this paper can prove useful in other settings and how this has been addressed in the paper.

3.1 Overall Approach

Based on the objective of this report, the method and the structure used to investigate the problems will follow a practical approach. This method is used to answer the overall research objective by introducing ideas with supporting arguments. The report is structured according to the three sub-approaches mentioned in section 1.2.2 with the objective to connect the theoretical base with findings to fit the result in an appropriate context.

3.2 Data Collection

In the following section the data used to answer the different research questions is presented and explained. As previously stated, the objectives of this paper are to complement existing research with deeper insights in how activist owners can create value through ESG-engagements and to provide insights and quantitative estimations of how such engagements can be conducted by Schroders with the most success.

The research questions are built upon the foundation of a data collection process and the objective is to gather information that will create supporting arguments and fulfill the overall objective of the report. The market information and data needed to answer the questions is based on both first- and secondary resources. By having the privilege of being provided with first hand resource data as well

as gathering information through secondary sources gives access to a wider range of information in a time efficient manner.

3.2.1 ESG Engagements

The ESG-engagement data is retrieved directly from Schroders and their internal information system. This way, it can be considered to be of high quality and will be of great relevance for the purpose of this report. Schroders have taken many different engagement actions over the years and showed a major commitment to responsible investment. Schroders became a UNPRI signatory in 2006 and in 2015, they managed over $ 44 billion of ethically-screened mandates (Schroders, 31 December 2015). According to Investment & Pensions Europe (2016), Schroders was ranked as the 31th largest Asset Manager in the world by AUM. Since the beginning of 2000 and until today (read 2017), Schroders have kept record of all their engagements and gathered detailed information about the results and outcomes that will be of great importance in making this report.

"Issues such as climate change, resource scarcity, population growth and corporate failure have put responsible investment at the forefront of investors' minds. We believe that companies with a strong environmental, social and governance ethos tend to deliver better results for our clients" - Peter Harrison, Group Chief Executive, Schroders

---

Schroders actively engages in dialogues that are held through in person meetings, emails, collaborative engagements and telephone conversations. If needed, their ownership rights are being exercised at shareholder meetings on behalf of internal and external clients (5,100 + in 2015). In addition to that, they are actively screening out ESG-irresponsible companies from some of their investment portfolios. The targeted companies are identified by applying different ESG-screening metrics and are chosen from Schroders current and prospective holdings. Schroders analysts examine companies' management of the environmental, social and governance challenges and the opportunities they face and summarizing in numerical rankings are relative to peers. Each quarter the ESG team at Schroders screens desk portfolios against third-party ESG ratings from specialist ESG research providers to identify holdings deemed to have poor ESG performance. These ratings are distributed to their various investment desks so each can assess the potential ESG risk. The ESG ratings are used as an input to company evaluation, risk assessment and finally investment conclusions.

The original dataset from Schroders consists of 4,628 engagement sequences (1,240 environmental, 1,222 social and 2,166 governance) for 1,033 public firms between 2000-2017. The focus of this research will be in the period between 2005-2015/2016 for two reasons. First, a time-series of eleven/twelve years has been deemed appropriate and it will provide both quantitative and qualitative accuracy in the tests. Secondly, according to information about the data given directly by Schroders, the earlier engagements and details surrounding them might not be as accurate and persist the same quality as more recent data, hence including earlier engagements could result in lower reliability. Applying these two reasons, starting off by removing data from all periods between 2000-2004 and 2016-2017, 1,506 observations are eliminated. Secondly, firms were excluded if it had an unrecognizable firm identification (-2) or if the requested data was not available for the firm (-66). This generates a total of 3,054 unique engagements in the period 2005-2015.

ESG-Component	**Engagements**
Environmental	956
Social	813
Governance	1285
Total	**3054**

Table 1: Number of engagements per ESG-component. This table presents the total number of engagements done over the time period 2005-2015 and how these are distributed across the ESG-components.

For the purpose of this research, there is no use in publishing the names or reveal the identification of these companies. By keeping it disclosed, there is no hindering Schroders in their process going forward within this field. However, what is of certain interest is the geographical location of these companies and what sector they belong to. Schroders have engaged with companies from a total of 39 different countries throughout this period. In general, Western Europe is well represented within the top ten (except for United States and China) and the United Kingdom in particular, who has more than twice the amount of engagements compared to France who came second.

Figure 4: Number of engagements per geographical area. This map shows the geographical dispersion of the number of engagements in period 2005-2015. Countries are categorized into regions defined by Morgan Stanley. The five regions are North America, South America, Europe; Middle east; Africa, Asia and Pacific.

All these companies are categorized using the Global Industry Classification Standard (GICS), developed by MSCI and Standard & Poor's (S&P). This standard includes 10 sector categories; material; financial; consumer discretionary; industrials; consumer staples; energy; utilities; healthcare; information technology and telecommunication services. Noticeable, is that the number of engagements is fairly balanced across all sectors, with a slight dominance from the materials sector.

Sector	Engagements
Consumer Discretionary	406
Consumer Staples	354
Energy	323
Financials	479
Health Care	156
Industrials	367
Information Technology	142
Materials	567
Telecommunication Services	88
Utilities	172
Total	**3054**

Table 2: Number of engagements per sector. This table presents the total number of engagements done over the time period 2005-2015 and how these are distributed across sectors.

There are two types of engagement approaches Schroders chooses to categorize as; Discussion and Objective. Discussion can be explained as a fact-finding approach where Schroders are in contact with a company to get a better understanding of their approach on ESG-issues. The second categorization, Objective, is when Schroders engage with the company and raises one or several requests for a specific change in their way of practicing ESG. To further specify the topics, Schroders have listed a series of subcategories (33 in total) which all fall in line under its natural category; Environmental, Social and Governance. Under environmental, the most frequently engaged topics are climate change, pollution and water, with climate change by far being the most common one as it represents almost 50% of all environmental engagements. Social is the smallest sub-sample and the topics with the most observations are human capital management, health & safety, customers and supply chain management. Within in governance, which is the largest category of engagements for Schroders, the engagements are divided more equally over a broader spectrum of topics. The ones standing out are corporate strategy, business integrity, remuneration, governance oversight and shareholders rights. These five together represents almost 80% of all governance engagements in this period.

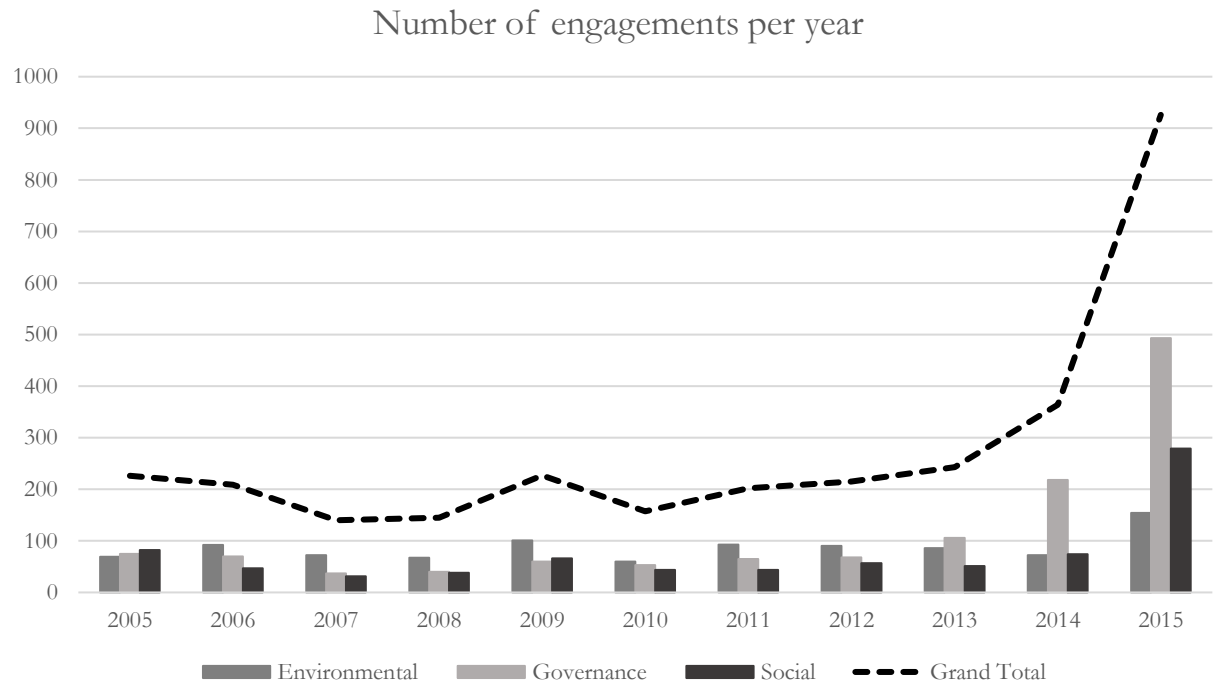


Figure 5: Number of engagements per year. This graph illustrates the number of engagements done each year. It is broken down into the specific ESG-component and the grand total for each year. The sum of all years is 3054

3.2.2 Financial Firm Specific Data

To obtain firm specific financial data, the historical corporate accounting figures have been downloaded from Thomson Reuters Datastream. Their database encompasses over 88,000 companies (including inactive ones), traded in more than 164 different exchanges in over 120 countries around the world. This represent more than 99% of the world's market cap. This capacity provides excellent comparability across companies and it is to be considered as a highly reliable and qualitative resource. The data has been downloaded and merged together with the data provided by Schroders through connecting ISIN-codes with SEDOL numbers and firm names. As mentioned earlier, when not able to retrieve complete information on the engaged company, it had to be excluded from the dataset. Worth mentioning is that before deciding on not to include these 66 companies, random controls were performed to find reasons why the information was not available. These random controls revealed quite a variety of reasons that spanned from mergers & acquisitions, delistings to bankruptcy.

3.2.3 Descriptive statistics[1]

This section will present graphs and tables that further explain the composition of the data used in the paper. The use of descriptive graphs and tables is not explicit for this section but throughout the report they will be included if there is a section discussing that specific area. The graphs and tables presented in this section is consequently the information that benefits the reader without the need to be complemented with extensive information. The information is included to allow the reader to grasp the scope of the data and to present insightful information that will be used in later parts of the study. Table four presents the most common topics for the engagements, for each ESG-component, with the number of engagements done on that specific topic. Corporate strategy and Climate change are the two most common topics

Topic	Engagements
Environmental	**956**
Climate Change	421
Pollution	273
Water	122
Biodiversity & Ecosystem Services	59
Others	81
Social	**813**
Human Capital Management	239
Health & Safety	142
Customers	127
Supply chain management	99
Human Rights	77
Labour Standards	70
Others	59
Governance	**1285**
Corporate Strategy	453
Business Integrity	351
Remuneration	168
Shareholder Rights	116
Board Structure	63
Others	134
Total	**3054**

Table 3: Most common engagement topics. This table presents the number of engagements within each ESG-component and the most common engagement topics. Topics with fewer than 50 total engagements have been classified as "Others".

[1] Please also see the case studies provided in appendix 9.6

for the engagements done and the group named "Others" contains topics with fewer than 50 engagements that have been classified together for the sake of availability to the reader.

ESG-component	Average number of engagements
Environmental	7.67
Social	7.72
Governance	6.69
Total	**7.28**

Table 4: Average number of engagements per ESG-component. This table shows the average number of engagements done for firms within each of the three ESG-components. Firms that have been engaged on governance topics show the lowest number of average engagements whereas the number is very similar between environmental and social.

Number of engagements	Number of firms
1	155
2	136
3	74
4	46
5	35
6	25
7	24
8	11
9	17
10	9
11	3
12	5
13	6
14	10
15	7
16	2
17	4
18	6
19	1
20-25	14
26-30	5
31-51	8
Total	**603**

Table 5: Unique engagements per firm This table presents the number of unique engagement topics for each firm. If a firm has been engaged on several topics at the same occasion, these are all reflected in these numbers. The right column shows how many firms are in each count category. The highest number of unique engagements for a firm is 51

Sector	Average number of engagements
Consumer Discretionary	5.45
Consumer Staples	7.43
Energy	9.80
Financials	10.00
Health Care	4.77
Industrials	4.63
Information Technology	3.35
Materials	7.13
Telecommunication Services	10.31
Utilities	8.27
Total	**7.28**

Table 6: Average number of engagements per sector. This table shows the average number of engagements done with firms in different sectors. The IT sector presents the lowest number of engagements whereas the Telecom sector has an average of 10.31 engagements per firm.

Table six describes the number of times a firm has been engaged and counts the number of firms in each category. It entails that the maximum number of engagements within any firm is 51. The fifth and seventh tables build from the same information but are categorized differently. They present how companies engaged have varying numbers of engagements in specific categories and show that across ESG-components there is little difference. This indicates that many firms are engaged on all three and that there is no component that goes before the others in the engagement sequences. The sector categorization however shows some differences and the table summarizes the idea that some sectors on average have a lower number of engagements per firm.

The last segment of this section focuses on how the topic type, i.e. the engagement approach varies between ESG-component, sectors and topics. Figure six shows how the number of engagements have varied between the years and is split into the two topic types. It also shows how the number of achieved objectives move in relation to these two. Interestingly, the number of achieved engagements in the first year is 94 and in the last year it is 103, showing a flat learning curve as the number of engagements have increased significantly during the time. Also, this figure visualizes the increases in engagements overall and how discussions have been the most represented type for the majority of the periods. The three tables eight, nine and ten all present the same information but are broken down into different categories. They illustrate how the total number of engagements are split between discussions and objectives and also includes a column with how many of the objective engagements that have been "Achieved". Table eight for example, shows that Social engagements are unique in the

sense that there have been more objective engagements than discussions and the environmental component has the highest success rate amongst its objectives. In table nine the information is broken down into the sectors and how the number of discussions and objectives vary for these. Lastly, table ten looks at the most common topics and how the engagements are split between these.

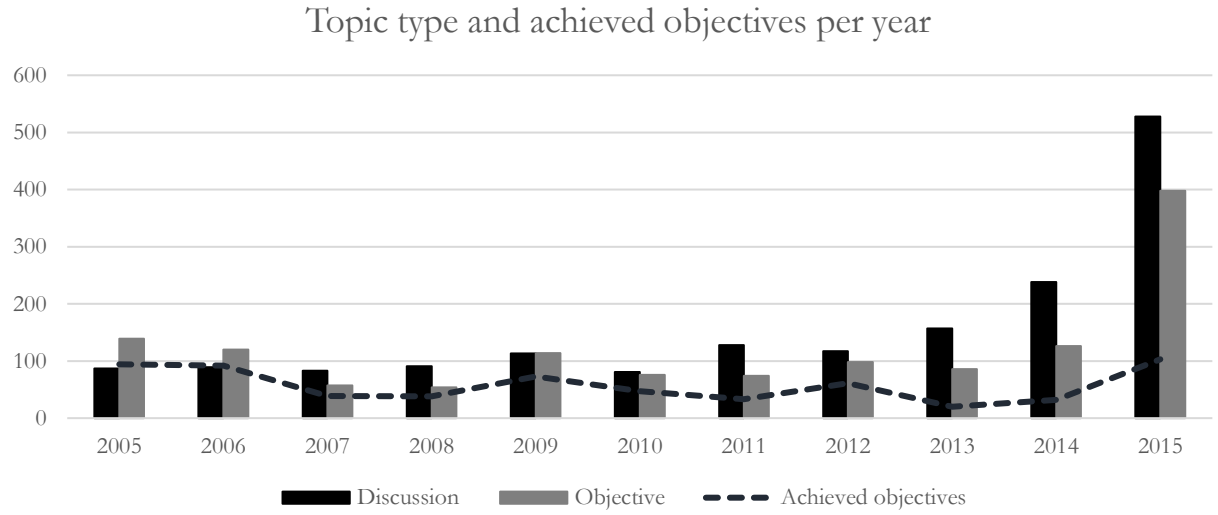


Figure 6: Topic type and achieved objectives per year. This graph illustrates how the number of engagements each year are divided between the two topic types Discussion and Objective. The number of achieved engagement objectives is also plotted.

ESG-Component	Discussion	Objective	Achieved objectives	Total
Environmental	608	348	51%	**956**
Social	382	431	45%	**813**
Governance	722	563	46%	**1285**
Total	**1712**	**1342**	**47%**	**3054**

Table 7: Engagements per topic type and ESG-component. This table presents how the number of engagements differ between topic type and ESG-components. The percentage of achieved objectives is derived from the number of achieved objectives divided by the number of objective engagements.

Sector	Discussion	Objective	Achieved objectives	Total
Consumer Discretionary	238	168	38%	**406**
Consumer Staples	211	143	48%	**354**
Energy	205	118	39%	**323**
Financials	211	268	44%	**479**
Health Care	85	71	46%	**156**
Industrials	209	158	47%	**367**
Information Technology	73	69	30%	**142**
Materials	318	249	55%	**567**
Telecommunication Services	56	32	78%	**88**
Utilities	106	66	68%	**172**
Total	**1712**	**1342**	**47%**	**3054**

Table 8: Engagements per topic type and sectors. This table presents how the number of engagements differ between topic type and different sectors. The percentage of achieved objectives is derived from the number of achieved objectives divided by the number of objective engagements.

Topics	Discussion	Objective	Achieved objectives	Total
Environmental				
Climate Change	270	151	48%	**421**
Pollution	178	95	66%	**273**
Water	76	46	57%	**122**
Biodiversity & Ecosystem Services	41	18	39%	**59**
Others	43	38	21%	**81**
Social				
Human Capital Management	84	155	44%	**239**
Health & Safety	56	86	45%	**142**
Customers	63	64	50%	**127**
Supply chain management	40	59	46%	**99**
Human Rights	52	25	48%	**77**
Labour Standards	52	18	72%	**70**
Others	35	24	17%	**59**
Governance				
Corporate Strategy	282	171	59%	**453**
Business Integrity	201	150	57%	**351**
Remuneration	126	42	36%	**168**
Shareholder Rights	11	105	29%	**116**
Board Structure	39	24	21%	**63**
Others	63	71	32%	**134**
Total	**1712**	**1342**	**47%**	**3054**

Table 9: Engagements per topic type and topics. This table presents how the number of engagements differ between topic type and different topics. The percentage of achieved objectives is derived from the number of achieved objectives divided by the number of objective engagements.

3.3 Reliability

In addressing the reliability of a study, the researcher assesses the potential of conducting another study with the same quantitative population, purpose and method and the ability to yield the same or similar results, making them reliable. With a quantitative study, the requirements for reliability are seen as tougher, as it usually tries to generalize a certain problem (Svenning, 2003). Ensuring reliability gives the author the opportunity to, in a structured manner, ensure that the data- gathering, processing, testing, analyzing and interpretations are sound and can be easily followed. With the exception of the private data from Schroders, this paper is based upon data from Datastream and Worldscope, two well recognized and easily accessed databases. The data from Schroders was to the most part structured qualitatively, so that its inclusion in this paper depended on it being quantified. The approach taken to this, creating (for the exception of the objective result variable) binary variables based on this information, removes any subjective aspect of this process and thus allows for easy and accurate replication.

3.4 Validity

Even with a reliable method and approach to a study, it is not certain that the study measures what it is intended to measure and thus it can differ in its ability to depict reality. This is referred to as the study's validity and there are two separate aspects to this, the internal and external (Svenning, 2003). The internal validity evolves around the reasoning and connection with empirical research and findings and how these relate to the structure and method of the paper. Internal validity is high when all variables and arguments that form the ground of the paper can be deemed valid and when any modification, alteration or addition is well reasoned and intuitive (Svenning, 2003). The external validity instead addresses the ability to generalize the results of the study beyond the given sample employed (Svenning, 2003). The method chosen has taken the steps available to broaden the usefulness and applicability of the results and with a more explorational focus, this paper is not explicitly attempting to prove or disprove causal relations between certain variables, but rather to investigate whether theoretical reasoning aligns with quantitative test results.

3.5 Robustness validation

This section will address some measures of fit for the regression models to be used and what aspects are important to consider when assessing the quality of the model and its estimates. This is to no extent an exhaustive description and the aspects considered have been based on their accessibility, comparability and ease of understanding. The details provided here will to the most part also be

presented in the result tables where the reader, with support from these sections, can interpret the variations in the models.

3.5.1 The Direction of Causality

How to determine which variable should be dependent and which should be independent is not always obvious and can be of a big concern (Bryman & Bell, 2011). Seeing as this study is more explorative, there will be short reflections on how causality can be decided. However, an important distinction is that this is not the primary goal of this study. Instead focus will be on investigating whether or not there are relations amongst the engagement variables and other variables of interest and to reflect on why this might be and how it can ultimately affect the business of Schroders. In presenting a complementary view to what previous research has found as of today, the paper does add to the understanding of the causality in the relations studied as it attempts to measure the effects from two ways and reflect on which seems more plausible and well-based.

3.5.2 Multicollinearity

Multicollinearity is an issue that can challenge and prevent the estimation of coefficients within a model. There are two cases of multicollinearity; perfect and imperfect (Stock & Watson 2012). Perfect multicollinearity arises when a regressor is a perfect linear combination of other regressors within the same model. This means that the variables become perfectly correlated and the model will not be able to generate the desired coefficients. Although the consequences of it seem dreadful, it is easy to avoid this issue. One approach is to remove one of the variables from the model. Seeing as all the variation in the excluded variable is in a sense captured by the other variables, the estimations does not suffer and the multicollinearity is removed. A second option, which perhaps is more logically sound, is to either redefine any of the variables or adjust the model accordingly. This could lead to the second sort of multicollinearity, imperfect multicollinearity. As opposed to perfect multicollinearity this does not make the model infeasible, however it might affect the estimated coefficients. Imperfect multicollinearity is the case where two or more regressors are highly correlated in the sense that there is a linear function of one regressor that is highly correlated with a linear function of another regressor (Stock & Watson, 2012). In general, when this situation occurs, the coefficients of one or more regressors are likely to be imprecisely estimated and have a larger sampling variance. If the variables are included with the motivation that they reduce the omitted variable bias and are logically sound, imperfect multicollinearity can be accepted in a model. A control for this can be done by looking at the correlations between the variables applied in the models and search for strongly correlated variables. Further, when included as control variables, imprecise estimates are not a direct concern as

the coefficient estimates of those variables will not be further interpreted but rather exist to generate a precise estimate of the variable of interest. Looking at the correlation matrix[2], none of the variables used across the models show alarmingly high correlations meaning that using this set of control variables, in relation with the performance measures, should be considered plausible.

3.5.3 Regression R^2

The regression R^2 measures how much of the sample variance of Y_i that is explained or predicted by the regressors in total. If the regressors are able to fully explain the variation in Y_i, the R^2 will take a value of 1. When considering multiple regression models like the ones used in this paper, R^2 has the trait that it always increases when another regressor is added. This means that unless the additional variable added explains zero of the variation in the dependent variable, which is extremely uncommon, the R^2 will increase. The definition for the R^2 is as follows: (Stock & Watson, 2012)

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{SSR}{TSS} \quad \text{where}$$

$$\textit{The Explained Sum of Squares } (ESS) = \sum_{i=1}^{n} \left(\hat{Y} - \bar{Y}\right)^2 \text{ and}$$

$$\textit{The Total Sum of Squares } (TSS) = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

When assessing the fit of a model based on R^2 one should recognize that adding variables to achieve a higher R^2 is not necessarily making the model a better fit. The variables should be included on the basis on whether it allows the model to better estimate the effect of interest. In this paper, the control variables fulfill this purpose and they have been assessed individually and cohesively to add predictive power to the model. What is worth noting is that the R^2 will only be reported for model 2, a model with panel data, and that this model is per definition in the software adding variables for each cross section and year. This results in a very high number of variables and consequently very high R^2 estimates. In that sense the R^2 is of limited use but the same model has been run (unreported) without the added cross sectional and time variables and proven to have a much lower, yet satisfying R^2. Using this measure indicates that the specified model is to various extent able to explain the variation in the dependent variable and a high R^2 indicates that the estimates of the variables of interest are well estimated.

[2] Please refer to the correlation matrix in appendix 9.1

3.5.4 The Standard Error of the Regression (SER)

The SER complements the R^2 as it is a measure of the spread of the observations around its regression line and it is an estimator of the standard deviation of the regression error term, u_i. The inherent nature of the SER is that the estimate is denoted in the same unit as the dependent variable and in effect it is measuring the magnitude of a typical error, in the unit of interest. This is a favorable trait as it makes it easy to interpret and to assess whether the error is large or small. Such assessment is ultimately subjective as it relates to how tolerant one can be of errors and their magnitude, and large errors indicates that there are likely other specifications of the model that could perform better estimates. The SER is defined as follows: (Stock & Watson, 2012)

$$SER = s_{\hat{u}} = \sqrt{s_{\hat{u}}^2} \quad \text{where} \quad s_{\hat{u}}^2 = \frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2 = \frac{SSR}{n-k-2}$$

The SER will be presented along the appropriate models and should be interpreted as a measure of fit for the models. It could also be of value when comparing models with the same dependent variable but different specifications of the variable of interest to evaluate whether certain time periods are more efficient in reducing the noise in the models. A lower SER indicates that predictions about the value of the dependent variable will be more precise.

3.5.5 The Wald χ^2

To test whether a model is a good fit for the data and if the variables included do help in understanding the variation in the dependent variable, the Wald test can be used. This is a standard way to perform large-sample inference testing and is an accessible way of determining whether the model is significant or not. In this sense the Wald test estimates if at least one of the variables included in the model adds predictive power to the model. The definition of the Wald test is somewhat tedious, so the essence of it will be presented here and more efforts will be put on how to interpret the results in relation to the models. Put differently, the Wald test has the null hypothesis that $\beta = \beta_0$ and where using nonnull standard errors of $\hat{\beta}$ the test statistic z is calculated by:

$$z = \frac{(\hat{\beta} - \beta_0)}{SE}$$

This statistic has an approximate standard normal distribution and the two-sided alternative has a χ^2-distribution with 1 degree of freedom. The multivariate extension of the Wald test has test statistics:

$$W = (\hat{\beta} - \beta_0)' \left[\boldsymbol{cov}(\hat{\beta})\right]^{-1} (\hat{\beta} - \beta_0)$$

The prime on a vector or matrix denotes the transpose (Agresti, 2002). What this test adds to the models is a confirmation that the relationship estimated has some value. Generally, a lower Wald χ^2 the lower predictive power the model holds. Important to note is that when this estimate is applied to and analyzed across models, with the same specifications but different variables of interest, it might differ between models. The importance and the reason for including it in the tests in this paper is consequently to make sure that the model holds across a great majority of the model specifications but that a varying degree of predictive power is expected and hard to avoid. This is accepted due to the fact that the predictive power of the models is not of primary interest but rather to compare differences in a specific variable across models, making it important to have the same model specifications.

4. Statistical Methodology

This section will start with some detailed description and elaboration on the different variables that are included in order to test the hypotheses. These have been separated into dependent variables, or performance measures, and explanatory variables that are taken from the sample from Schroders and firm specific control variables. In order to understand the models and their implications, it is important to understand the underlying variables and their reason for inclusion as well as their shortcomings. Furthermore, this section will explain the various methods and models applied to the data set introduced, and how such approaches are helpful in testing the hypotheses, that will be presented simultaneously. Due to the many areas of interest and the relatively low frequency of tests on data similar to what is used in this paper, the statistical methodology will take numerous approaches. This is for two main reasons: The first one is that by investigating a spectrum of effects, it is possible for both Schroders and future research to narrow the scope and look at certain questions in more detail. Secondly, the approach is motivated by the notion that the research on ESG has not yet reached a level where information, data and literature are supporting highly detailed testing. This will become clear and addressed further in the model descriptions that follow.

4.1 Performance Measures

Performance measures are subjective measurements of how well a firm has used its assets from its primary mode of business to generate revenue. The different measures are used to provide a general measure of a firm's overall financial health over a given period of time and it can also be used for comparison between similar firms across the same industry or sector. Return On Assets (ROA), Tobin's Q, Sales per employee and Excess return the four variables that will be used as performance indicators across the models. They are all to some extent standardized and well-known measures but will be addressed to some detail here to point at their contribution to the paper.

4.1.1 Objective result

The Objective result has some various traits to it. The first important distinction is that for items with Topic Type "Discussion" there is only one given Objective result, namely "No further change required". This makes these observations less informative for some reasons. Firstly, there is no difference between engagements and thus it is not possible to study the effects of them at a high level. Secondly, this type of engagements does not require any counteraction from the target firm which removes any ability to measure or interpret how the engagement has been received by the target.

Ultimately, this classifies all such engagements as successful or achieved, even if the only action is taken by Schroders. Due to this, these items (n=1712 in whole sample 2005-2015) are excluded when Objective result is used as the dependent variable, ultimately generating a sample of 1341 engagements. Now having only Objective types of engagements, the data can be further broken down into their levels of success. There is a clear hierarchy in the categories of the results. The first one is "No change" which simply means that there have been no advancements by the target firm towards the suggested change for 12 months. The next level of advancement is "Some change" which indicates that there has been a response from the target but that this is not yet is close to completing the change or implementation process. The third category "Almost" is an indication that the most part of the suggested activity has been carried through but that there are still some aspects left to be addressed. The final category is "Achieved" and recognizes that the target has fulfilled the items agreed upon and that the engagement has been successfully completed. The result "No further change required" is also plausible for objective engagements and is interpreted somewhat differently in these cases. It indicates that the engaged company has been unable to succeed with the suggested change before some outside factor made the objective irrelevant. Such cases could be changes in regulation, mergers or acquisitions or if the engagement has been active for more than 5 years without a change. These observations thus end up somewhere between "No change" and "Almost" however it is impossible to concretize further. This offers two potential options as to how to handle this information in the data set used in the models. Either the engagements are removed and excluded, or they are classified into one and the same existing category. Deciding to count them as "No change" is based on the knowledge about the engagements specifically and the fact that given a time interval, the change initiative was not accomplished and thus not successful as it did not result in the outlined activities. There are 68 such occasions in the data and when coded for the models, these have been given a value of 0, together with the engagements with "No change". More detail on this is provided in the model specification below.

4.1.2 Return On Assets

As a performance measure, ROA measures the amount of profit a company generates as percentage of the book value of its total assets. It illustrates how well management is employing the company's total assets to make a profit. It is calculated as the ratio of its net income in a given period to the total value of its assets. The profit percentage of assets varies between sectors. For this reason, it can be more effective to compare a company's ROA to that of other peers in the same sector or against its own ROA from previous periods. In general, the higher ROA the better, and a falling ROA is almost

always a problem. As mentioned, the ratio is more useful in some industries than in others, partly because how much money this particular company has tied up in assets will depend on a given sector. However, regardless of different industries, the measure will provide a bigger picture of the overall return for the firms in this project. ROA is by no means a perfect measure, but it is a highly effective, broadly available financial measure to assess a company's performance. The fundamentals of business performance are being captured in a holistic way, through looking at both income statement performance and the assets required to run a business. Other metrics could have been used, such as the return on equity, but many of them are vulnerable to financial engineering, especially through debt leverage, which can obscure the fundamentals of a business. One of the strongest arguments for the use of ROA is how persistent it is to the kind of short-term gaming that can occur on income statements since many assets, such as intangibles, property, plant and equipment, involves more long-term decisions that are more difficult to tamper with in the short term.

4.1.3 Tobin's Q

Another measure of performance, Tobin's Q, calculates the ratio of the market value of a firm's assets (as measured by the market value of its outstanding stock and debt) to the replacement costs of the firm's assets (Tobin, 1969).

Bharadwaj, Bharadwaj & Konsynski (1999), mentions in their research how widely used Tobin's Q has been in previous research to explain a wide variety of phenomenon such as; an alternate measure for business performance, a predictor of profitable investment opportunities and a measure of the value of a firm's intangible assets. Being a financial market measure, Tobin's Q has many attractive aspects as it is not only based on theoretical and empirical foundations of the efficient market hypotheses, but also addresses a growing concern over the limitations of accounting measures of performance. They are also pointing out the fact that market measures, which have a long tradition in the corporate finance literature, are presumed to have some of the following advantages; stock prices represent the only direct measure of stockholder value, stock prices fully reflect all available aspects of performance, stock prices are not only reported objectively but are also easily available for publicly traded firms and stock prices can "see through" managers attempts trying to manipulate reported accounting measures. Consistent with the forward-looking nature of the capital markets, the Tobin's Q ratio measures the market power from both existing assets and the firm's future growth potential. Tobin's Q has been a popular choice and used extensively when measuring the value of a firm's

intangible assets. Measures of intangible values are based on the assumption that the long-run equilibrium market value of a firm must be equal to the replacement value of its assets.

4.1.4 Ln (Sales per employee)

When taking a stakeholder perspective of the firm, one actively chooses to recognize value creation for other stakeholders rather than solely the shareholders. In doing so, there are some internal and external stakeholders that might react and contribute to this value creation differently. Externally, a stakeholder view can be seen through improved customer loyalty or lower price elasticity as customers put a higher value on the firm's operations and products. Internally, employees can be considered to be a heavily influenced stakeholder that also has potential to influence and alter the activities of the firm. The efficiency at which the employees of a firm work is highly reflected not only in revenues but also in profit margins and ultimately the performance of a firm. Having productive employees is a key driver of profitability and when management are aligning the firm's operations with the interests and ideals of its employees, this might positively influence their motivation, retention and aspirations to further improve the firm. The Sales per employee variable is able to measure how much revenue is generated by the average employee and if also controlling for sales growth itself, it provides useful insights into how changes in revenues can stem from each employee generating more (Dimson et al., 2015). The attractiveness of this measure is dual. Firstly, it can estimate a tendency with employees to alter their performance based on the activities of the firm and thus not only potentially strengthening the benefits of ESG-initiatives, but also affirming aspects of the stakeholder view of the firm. Secondly, it complements the other performance measures as it focus explicitly on the employee contribution (assuming that changes in sales is controlled for, which is the case in this study) and thus provides an internal viewpoint as contrasted by the other measures also considering external aspects in the measures. Flammer (2015) recognizes that the adoption of close call shareholder votes on CSR activities has the potential to improve labor productivity and that one of the channels through which this is done is that such proposals increase job satisfaction and that it enables the firm to reach customers that are responsive to such initiatives with lesser effort.

The influence of employees not only as a source of revenue creation and a resource but also as a risk reduction tool, has further emphasized the importance to create a mutually beneficial relation between firm and employees. It can be argued that a stronger relationship with employees not only can *create* superior performance but also *preserve* it and thus create a competitive advantage (Godfrey et al., 2009). Recognizing the multiple ways that ESG-initiatives of different sorts can enhance employee

satisfaction, efficiency, relations and other productivity enhancing aspects, investigating whether the engagements in the data set have been able to accentuate such is of great interest (Galema, Plantinga & Scholtens, 2008).

4.1.5 Excess return

Looking at accounting measures provides a first step in understanding how a firm can be affected by altering their business operations to incorporate more ESG-aspects. How this is accepted amongst the investors is ultimately decided by the changes in the share price and its effect on future dividends and/or share repurchases. Taking the investor's perspective, it can be seen as irrelevant if the firm is improving on ESG-issues if it is not something valued by the market or reflected in payouts. There are numerous ways to estimate the over- or underperformance of individual stocks and portfolios, also within the field of SRI. Many of these use sophisticated models assessing a risk adjusted return in a portfolio setting where a benchmark of either "bad" companies or comparable index is used, for some examples addressing this in a somewhat similar setting please see reference (Galema et al., 2008; Smith, 1996; Barnett & Salomon, 2006; Becht, Franks, Mayer & Rossi, 2010). In employing several multi-factor models to the analysis of abnormal returns, the shortcomings of any individual model can be complemented by the application of another model. However, such an analysis is not the primary objective or approach of this thesis. Here, it has been deemed to better fit the purpose and design of the study to use a more standardized model. Thus the Capital Asset Pricing Model (CAPM) has been used to establish expected returns for each firm-year. Though criticized in literature for relying on assumptions that does not hold empirically, primarily the assumption that investors are rational, it remains one of the most applied methods for estimating expected returns (Pratt & Grabowski, 2014). The model outlines the following relationship for expected returns:

$$E(r_i) = rf + \beta\,(r_{mkt} - rf)$$

Where; $E(r_i)$ is the expected return for firm i; rf is the risk free rate; β is the estimate of a firm's systematic risk; r_{mkt} is the market return.

This has been estimated for all firms and all years between 2005 and 2016. The risk free rate and the market risk premium has been retrieved from Kenneth R French's (2017) digital archive. The risk free estimate is the one month US T-bill rate and the market return is created using a global value weighted market portfolio that takes into consideration the dividends and capital gains but is not annually compounded. These estimates provide a coherent estimation of the risk free rate and market return

as they have been brought forward by Mr. French for the purposes of assessing expected returns in different settings, with different models. The fact that the market return is based on a value-weighted global index is appropriate and matches the data set used in this report. Recognizing that the engagement frequency is skewed for larger and more economically powerful countries such as the UK, the US and neighboring countries, using a market return estimate that has a similar distribution improves the estimates. The beta estimate is retrieved from Datastream and is derived *"by performing a least squares regression between adjusted prices of the stock and the corresponding Datastream market index. The historic beta so derived is then adjusted using Bayesian techniques to predict the probable behavior of the stock price on the basis that any extreme behavior in the past is likely to average out in the future."* (Datastream definition). To arrive at an excess return for each firm year, the actual return realized is needed. This has been estimated through Datastream by taking the percentage change in price, without reinvestment of dividends over the last 12 months.

	Objective result		**ROA**		**Tobin's Q**		**Ln Sales per employee**		**Excess return (%)**	
ESG-component	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**
Environment	1.85	1.32	4.58%	9.64%	1.11	0.97	6.41	1.61	1.77	59.96
Social	1.61	1.40	4.89%	9.65%	1.26	1.19	6.01	1.59	-2.96	29.10
Governance	1.77	1.30	3.88%	9.69%	1.14	1.19	6.16	1.63	-4.25	37.77
Sector										
Consumer Discretionary	1.64	1.28	7.17%	8.33%	1.49	1.34	5.43	1.61	-4.29	28.90
Consumer Staples	2.02	1.14	6.25%	5.49%	1.38	0.68	5.76	1.33	-0.81	21.26
Energy	1.47	1.37	2.93%	14.65%	0.94	0.80	7.58	1.11	-5.66	29.36
Financials	1.41	1.46	0.61%	9.41%	0.50	0.78	6.61	1.41	-7.98	24.13
Health Care	1.63	1.41	8.07%	13.38%	2.34	2.38	5.84	1.08	-2.83	25.30
Industrials	1.77	1.32	3.83%	5.84%	1.11	1.18	5.49	1.53	-2.42	28.92
Information Technologies	1.16	1.34	6.31%	11.72%	1.69	1.17	6.01	1.60	6.97	36.33
Materials	1.91	1.34	4.66%	9.12%	1.19	0.72	6.30	1.51	3.00	79.69
Telecom Services	2.50	1.02	1.53%	8.94%	0.96	0.43	6.30	1.59	-7.83	17.43
Utilities	2.50	0.90	3.80%	5.98%	0.82	0.41	6.73	2.18	1.38	30.16
Full sample	**1.72**	**1.35**	**4.37%**	**9.67%**	**1.16**	**1.12**	**6.20**	**1.62**	**-2.24**	**39.72**

Table 10: Average values & standard deviations of all performance measures. This table summarizes the average values and standard deviations of all the performance measure variables used in the models. The objective result has been given the following values based on the classification in the data. 0 = No change, 1 = Some change, 2 = Almost, 3 = Achieved. (For further description please see section 4.3.1.)

4.2 Control Variables

The following section will present the variables to be used as control variables in the models. They will only be introduced shortly but with references to previous literature that has applied these in much a similar manner. The variables chosen have been assessed to, intuitively, have a potential impact on the dependent variables to be estimated and thus might have a significant influence on the coefficients.

- **Firm size:** Defined as Ln(Total assets)[3]
- **Firm risk:** Defined as Debt/Equity ratio[4]
- **Liquidity:** Defined as Cash/Assets[5]
- **Value of intangible assets:** Defined as Market/Book ratio[6]
- **Growth:** Defined as $(Sales_{t+1} – Sales_t)/Sales_t$ [7]
- **R&D intensity:** Defined as R&D/Assets[8]
- **Capital investments:** Defined as Capital expenditures/Assets[9]

4.3 Model definitions

As mentioned in the outlining of this chapter, the models chosen to test the data presented will be elaborated upon in this section.

4.3.1 Model 1 - Objective results

The goal of this model is to investigate if there are certain characteristics of an engagement that are driving the objective results and by doing so improves the probability of the engagement being achieved. To test this, the information categorized as the objective result of an engagement will be quantified and used as the dependent variable in the model. As mentioned in section 3.3 there are four different values that the objective result can take is "No change/No further change required", "Some change", "Almost", and "Achieved". This type of data is labeled ordered response data and it arises when a variable is defined by mutually exclusive qualitative categories that has an inherent ordering, rather than defined as continuous numbers. This lack of natural numerical values makes regular OLS inappropriate (Stock & Watson, 2012). One major reason for this is clear when interpreting linear regression coefficients. They estimate the expected unit change in a dependent variable given a one-unit change in the independent variable. This interpretation becomes inappropriate when dealing with ordinal variables seeing as linear regression assumes that the difference between two values of the

[3] (Margolis et al., 2007; Barnett & Salomon, 2006; Bebchuk, Brav & Jiang, 2015; Bhojraj & Sengupta, 2003; Brammer & Millington, 2008; Dhaliwal et al., 2014; Edmans et al., 2011; Godfrey et al., 2009; Jo, Kim & Park, 2015; Jo & Harjoto, 2011; Smith, 1996; Siew, Balatbat & Carmichael, 2016; Grewal, Serafeim & Yoon, 2016; McWilliams & Siegel, 2000)
[4] (Bhojraj & Sengupta, 2003; Grewal et al., 2016; Brammer & Millington, 2008; Dhaliwal et al., 2014; Edmans et al, 2013; Jo & Harjoto, 2011; McWilliams & Siegel 2000; Smith, 1996; Margolis et al., 2007; Siew et al., 2016).
[5] (Brammer and Millington, 2008).
[6] (Godfrey et al., 2009; Dhaliwal et al., 2014; Edmans et al., 2013; Jo, Kim & Park, 2015).
[7] (Edmans et al., 2013; Jo, Kim & Park, 2015; Jo & Harjoto, 2011).
[8] (McWilliams and Siegel, 2000; Jo & Harjoto, 2011; Grewal et al., 2016).
[9] (Grewal et al., 2016; Jo & Harjoto, 2011; Jo, Kim & Park, 2015).

variable is the same for all intervals of that variable. This assumption is more difficult to justify when considering categories rather than continuous values and with the ordered probit model, this assumption is relaxed in the model specification and the coefficients can be estimated efficiently without this having to be the case (Daykin & Moffatt, 2002). A further discussion around these criteria is provided below.

In order to make use of categorical data that is not quantified, the use of logit or probit models is often applied (Stock & Watson, 2012). To get a better understanding of the ordered probit model used in this paper, first consider a logit or probit model where the dependent variable has a binary value of 0 or 1. A logit or probit model estimates the probability that the dependent variable takes a value =1. The difference between the logit and probit model is that a logit model assumes a cumulative standard logistic distribution (F) whereas a probit model assumes a cumulative standard normal distribution ($\boldsymbol{\phi}$). The probit regression model with a binary dependent variable is specified as follows:

$$Pr(Y = 1|X_1, X_2 \ldots, X_k) = \Phi(\beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots, \beta_k X_k)$$

The interpretation of β_1 in this model is that the coefficient represents the change in the z-value arising from a unit change in X_1, holding constant the other regressors. Although the effect of X on the z-value (in the cumulative standard normal distribution) is linear, its effect on the probability is nonlinear. The probit coefficients produced have been estimated using the maximum likelihood method that ensures efficient estimators (Stock & Watson, 2012). Although an ordered probit model has been used, the implications of the model remain the same.

	Ln Assets		Debt to Equity ratio (%)		Cash / Assets		Market to Book ratio		Sales growth		R&D / Assets		CapEx/Assets	
ESG-component	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**	**Average**	**S.dev**
Environment	17.30	2.08	137.45	595.30	0.06	0.07	2.78	10.07	6.28%	29.67%	0.01	0.03	0.06	0.04
Social	17.20	2.27	174.09	393.80	0.06	0.08	4.35	21.92	5.73%	25.27%	0.01	0.04	0.05	0.04
Governance	17.12	2.40	144.14	324.41	0.06	0.07	3.34	17.10	6.22%	34.50%	0.01	0.03	0.04	0.04
Sector														
Consumer Discretionary	16.12	2.15	115.47	383.34	0.09	0.09	8.05	38.88	8.06%	31.24%	0.01	0.02	0.05	0.04
Consumer Staples	17.03	1.61	47.93	374.85	0.06	0.05	1.95	14.75	4.14%	12.29%	0.01	0.01	0.05	0.03
Energy	18.05	2.22	49.22	53.88	0.05	0.05	1.94	1.67	1.16%	28.66%	0.00	0.01	0.10	0.05
Financials	19.12	2.34	441.44	500.27	0.03	0.06	1.39	1.29	4.88%	37.14%	0.00	0.00	0.01	0.02
Health Care	16.80	1.57	56.15	52.03	0.07	0.08	4.26	4.06	7.56%	14.34%	0.06	0.06	0.04	0.02
Industrials	16.04	2.02	140.00	825.54	0.08	0.07	4.61	14.98	4.45%	16.49%	0.02	0.03	0.04	0.03
Information Technologies	16.37	2.78	48.03	67.64	0.14	0.13	4.84	21.77	13.30%	23.06%	0.07	0.07	0.04	0.04
Materials	16.81	1.75	69.25	103.38	0.05	0.05	2.29	1.90	4.88%	27.66%	0.01	0.02	0.07	0.04
Telecom Services	18.25	1.42	153.13	228.97	0.04	0.04	3.94	7.12	6.15%	22.42%	0.00	0.01	0.07	0.04
Utilities	17.34	1.94	274.81	551.24	0.05	0.08	3.13	2.57	18.60%	67.73%	0.00	0.00	0.06	0.02
Full sample	**17.20**	**2.27**	**150.02**	**443.37**	**0.06**	**0.07**	**3.43**	**16.82**	**6.11%**	**30.75%**	**0.01**	**0.03**	**0.05**	**0.04**

Table 11: Average values & standard deviations on all control variables. This table summarizes the average values and standard deviations of all the control variables used in the models. Ln Assets is defined as the natural logarithm of the total assets of the firm. Debt to Equity ratio is the percentage of the firm's total amount of short- and long-term debt over stockholder's total common equity. The Cash/Assets ratio is retrieved by taking the total cash available over total assets. The Market to Book ratio is calculated by the price per share divided by the book value per share. The sales growth is simply the difference in sales from year t to year t+1 divided by the sales in year t. Taking the spending on R&D over the total assets of the firm has generated the R&D/Assets variable. Similarly, the spending on capital expenditures has been divided by total assets to generate the CapEx/Assets variable.

An ordered model is appropriate when the dependent variable have more than two possible values where one value is truly higher than another. This is the case for how successful an engagement has been and in this model, only objective engagements are considered as they attempt to trigger a tangible change and thus also a counteraction that is informative and interesting. In the data there are four potential results that have been coded and ranked in the following way:

- No Change/No further change required = 0
- Some Change = 1
- Almost = 2
- Achieved = 3

In this approach, the numbers do not mean much but they are rather a method of ranking the results for the statistical software. The way these works, in contrast to a binary probit model, is that we cannot truly estimate the percentage chance of being in a certain category, however we can interpret the independent variables as positively, negatively or neutrally affecting the probability of being in the highest category, achieved. More specifically, the model is designed following the approach by Daykin and Moffatt (2002) [10];

Let i index each engagement where $i = 1\ldots.n$ and n is the number of engagements in the sample. Y_i represents the level of success in each engagement and as noted above, this range is set to 0-3. This entails that Y_i^* $(-\infty < Y_i^* < +\infty)$ represents the underlying latent variable related to engagement *i's* propensity to achieve the suggested change. X_i is defined as the variable of interest and Z_i as a vector of firm specific characteristics that function as control variables. This model rests on the assumption that Y_i^* depends linearly on X_i according to:

$$Y_i^* = X_i\beta_i + Z_i\beta + u_i \ , \qquad Where\ i = 1\ldots, n\ and\ u_i: N(0,1)$$

The vector of β- parameters will ultimately contribute to the interpretation in similar ways as slope parameters in a linear regression. In this setting, Y^* is unobservable but the relation between Y^* and Y is:

[10] Please refer to the appendix 9.2 for further specification of model 1

$$Y = 0\ if\ -\infty < Y^* < k_1$$
$$Y = 1\ if\ k_1 < Y^* < k_2$$
$$Y = 2\ if\ k_2 < Y^* < k_3$$
$$Y = 3\ if\ k_3 < Y^* < +\infty$$

The parameters k_1, k_2 and k_3 can be referred to as cut points or threshold parameters. The cut points and the β-coefficients in the vector will not be interpreted broadly across the models, but it is rather the coefficient of X_i that will be elaborated upon. Figure 7 provides an illustration of the density function of Y^* and $x'\beta$ in the figure represents β_i and depends on the explanatory variable of interest. Following this dependence, the whole distribution shifts when there is a change in X_i and ultimately the model returns the probability that an engagement that meets the criteria for X_i reaches the highest value for Y, in this case 3 signaling an achieved change. The model is further specified as a log-likelihood function and the details of this can be found in appendix 9.2 as the description does not add further interpretational value. This last step allows the model to generate maximum likelihood estimates and asymptotic standard errors.

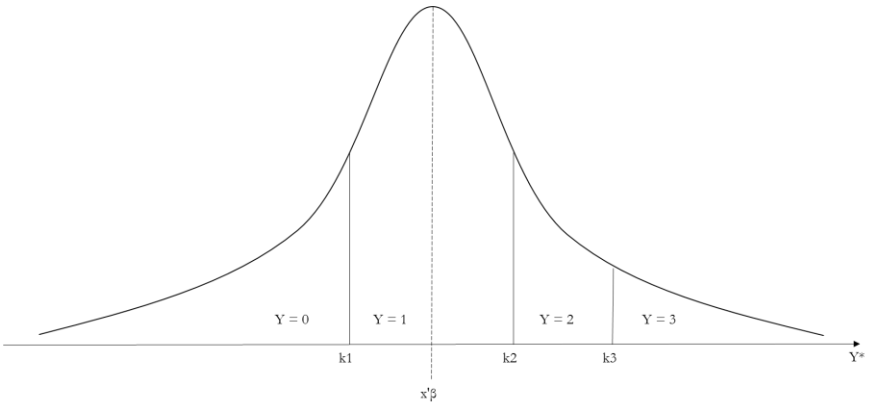


Figure 7: Density function of Y*. Adopted from Daykin & Moffatt (2002)

As recognized previously, one important distinction in this model is whether the "steps" can be assumed to be of equal magnitude. Stated differently; *"The values of an ordinal variable can be put into a unique order, but the distance between values cannot be quantified."* This means; is the difference between 0 and 1 equal to the difference between 2 and 3 in terms of how the variables are categorized? If a standard linear regression model had been applied, this would have been assumed to be true and thus the estimated coefficients had not necessarily been efficient (Daykin & Moffatt, 2002). Noting that the variable is defined, interpreted and applied by Schroders internally, and that the categorization ultimately relies on the subjective decision of an employee (or perhaps several), such an assumption seems to be questionable. A brief interpretation of the categories supports this as it can easily be argued that for a specific change to go from "No change" (=0) to "Some change" (=1) requires significantly smaller efforts than moving between subsequent categories. This varies depending on the

nature of the change and the complexity it captures, however there is no logical reasoning supporting equal difference between categories. The use of an ordered probit model is thus supported. Contrasting this, yet affirming the usefulness of using such a model. Pasta (2009) argues, that the distance between continuous variables does not per definition have a linear relation either. Consequently, the assumption of linearity between discrete values in a variable definition is highly dependent on the preferences of the researcher. This arguing tends to be supported by practitioners and the differences in estimates are often times insignificant (Williams, 2006). The model will test the whole sample to look at the aggregate effects for ESG but it will also be broken into three sub-samples, defined by the ESG-components to investigate more detailed relations. The results from this will be interpreted and further explained in chapter five.

4.3.1.1 Hypotheses Model 1

This model will test the following hypotheses, with the null hypotheses being that no difference is observed between the characteristics tested. The alternative hypothesis will be accepted if the probability of one type of engagement being "achieved" is significantly different from the probability that another type of engagement is achieved.

(1) H_0= There is no difference in the probability of achieving an engagement between ESG-components

    H_1= There is a significant difference in the probability of achieving an engagement between ESG-components

(2) H_0= There is no difference in the probability of achieving an engagement between engagement topics

    H_1= There is a significant difference in the probability of achieving an engagement between engagement topics

(3) H_0= There is no difference in the probability of achieving an engagement depending on the collaborative and repetitive nature of the engagement

    H_1= There is a significant difference in the probability of achieving an engagement depending on the collaborative and repetitive nature of the engagement

(4) H_0= There is no difference in the probability of achieving an engagement depending on the geographic region

    H_1= There is a significant difference in the probability of achieving an engagement depending on the geographic region

(5) H_0= There is no difference in the probability of achieving an engagement depending on the sector

H_1= There is a significant difference in the probability of achieving an engagement depending on the sector

(6) H_0= There is no difference in the probability of achieving an engagement depending on ROA

H_1= There is a significant difference in the probability of achieving an engagement depending on ROA

(7) H_0= There is no difference in the probability of achieving an engagement depending on Tobin's Q

H_1=There is a significant difference in the probability of achieving an engagement depending on Tobin's Q

(8) H_0= There is no difference in the probability of achieving an engagement depending on the efficiency, defined as Sales per employee

H_1= There is a significant difference in the probability of achieving an engagement depending on efficiency, defined as Sales per employee

(9) H_0= There is no difference in the probability of achieving an engagement depending on firm specific characteristics

H_1= There is a significant difference in the probability of achieving an engagement depending on firm specific characteristics

4.3.2 Model 2a - Corporate Financial Performance

Having investigated whether certain characteristics make engagements more likely to succeed, model 2 will establish if there are any relations between certain types of engagements and CFP-measures. The variables chosen to represent CFP are ROA, Tobin's Q and Sales per employee. They have all been presented above in section 4.1 and their inclusion is based on both on their individual contribution but also on their joint contribution. To some extent, these three all reflect on the internal efficiency and profitability with which a firm operates and to complement the previous tests with such perspective is important to address the research questions and to further explore any existing relations. However, they also differ in many senses which is why it is important to include them all and interpret their results and implications individually. Model 2a will investigate this through a continuous approach and model 2b will test the same variables but in a binary way, seeing if there is an improvement or not. Together the two will provide a broad, yet non-exhaustive, perspective on the operational value of engagements. First, the dependent variable Y_{it} is defined as a continuous variable,

representing the CFP-measure for firm *i* the year of engagement, time *t*. The interpretation of the estimated coefficients is straightforward; what effect does having a 1 instead for 0 (meaning that the firm has been engaged in the specific ESG-component or achieved an engagement that same year) have on Y. The fact that operational changes might not give an immediate effect but rather is something with a longer time horizon motivates the inclusion of two years lag on the dependent variables. Thus, the dependent variables will be Y_{it}, Y_{it+1} and Y_{it+2}. These test are conducted in order to estimate the potential magnitude of the effects that the engagements can have on the operational performance measure and the coefficients will be interpreted as absolute values, meaning that a coefficient β of 0,01 proves a one percent increase in the dependent variable if the engagement variable $X_1 = 1$ rather than 0.

This model, model 2a, follows a standard panel data model with two way fixed effects. The number of cross sections, i.e. unique firms, denoted as *i*,, in the sample varies depending on the model definition but all have been engaged by Schroders at least once during the time period 2005-2015 with the date of the engagement as the time defining variable, denoted *t*. The panel data model relies on the fact that you have n different entities that are observed during a time period T. Applying this to the data set used, the 11 years of observations, where a firm might or might not have been engaged, it means that firm *i* has only been engaged in a few years whereas another firm might have been engaged in other years. Thus, the values for *t* will not be the same for all firms. This makes the data unbalanced. Unbalanced data is defined as *"A panel that has some missing data for at least one time period for at least one entity…"* (Stock & Watson, 2012) and the years without an engagement for firm *i* means that the variables X_i and Y_i will not be observed in that year, creating the unbalance. This does not pose any restrictions on the testing of such a model, but rather requires more of the researcher to compose the data set in an appropriate manner. This will be addressed more in detail below. The two way fixed effects model applied is defined as follows; (Stock & Watson, 2012)

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_{it} + \alpha_i + \lambda_t + u_{it}$$

Where; Y_{it} is the defined CFP-measure for firm *i* at time *t*; β_0 is the intercept; β_1 is the estimated coefficient of variable X for firm *i* at time *t*; $\beta_2 Z_{it}$ represents a vector of the control variables included and their respective estimated coefficients; α_i is the entity fixed effects; λ_t is the time fixed effect and u_{it} is the error term.

The entity and time fixed effects are the primary benefit of using a panel data model and in effect they control for omitted variable bias and makes the estimations more precise. The entity fixed effect controls for unobserved variables that varies from one firm to the next but does not change significantly over time. This could for example be cultural aspects affecting not only the operations of the firms but also the attitudes towards ESG or geographical aspects. In the model above $\alpha_1 \ldots, \alpha_n$ can be seen as the effect of being in firm *i* and though the population regression line β_1 is the same for all firms, the intercept of the population regression line varies from firm to firm, a distinction enabled by including α_i. This way, omitted variables are captured and reflected in the estimation and so it is possible to estimate β_1 holding constant these out-of-model variables (Stock and Watson, 2015). Similarly, the time fixed effects control for variables that are constant across entities but evolve over time. In much the same way as the entity fixed effects creates an intercept for each firm, the time fixed effects includes an individual intercept for each time period in the regression model (Stock & Watson, 2012). This year unique intercept can be interpreted as the "effect" on Y that year *t* might have and in much the same way controls for omitted variables and makes the estimates of β_1 even more accurate. Using the combination of these, as stated in the model definition above, efficiently eliminates omitted variable bias that can arise both from unobserved variables that are constant over time and from unobserved variables that are constant across firms.

In a model specification based on panel data, the regression error might correlate with itself over time within one entity (Stock & Watson, 2012). One example could be the case for firm specific variables that are not included in the model and does not change from one year to another, such as board composition or which stock exchange one is listed at, but that might influence $Y_{it.}$ This correlation over time does not affect the estimations of the coefficients and does not introduce biases into those, however it affects the variance of the fixed effects estimator and thus the standard errors (Stock & Watson, 2012). Such autocorrelation is a common feature of time series data as what happens or is true one year is reasonably correlated to what happens or is true in the next year. If this is the case, using heteroscedasticity-robust standard error formulas, such as for cross-sectional regression models, the regression errors will be invalid. Consequently, heteroscedasticity-and autocorrelation-consistent (HAC) standard errors are applied to the tests to ensure that any regression assumptions are not violated. In defining the model, the way as has been presented, the coefficient estimates are estimated capturing as much information as possible, controlling for outside effects and variations and creates valid confidence intervals (Stock & Watson, 2012).

Having defined the model in more detail, fitting the data set into a structure that allows such tests is of major interest as the structural efforts are heavily rewarded by a model that significantly outperform other model definitions in terms of efficiency and consistency. Recognizing that any firm *i* can have several engagements in the same year *t*, a panel data model will not function with the data structured in such a manner. Thus the data has been re-structured and categorized following the three ESG components and instead of having one data set, three different sets have been used. In each specific E, S or G set, when a firm has several observations in the same year, the observation with the corresponding engagement component has been prioritized and kept, where the remaining firm observations in the same year has been removed. This creates three sub-sets that all prioritize the engagements done in their respective ESG-component. Further, a binary variable taking the value of 1 if a given engagement has been achieved or not and zero otherwise has been created. This variable will be tested independently and does not include the E, S or G specific prioritization as described but is rather defined to capture the eventual effect that achieving an objective has, allowing for comparison across the sub-samples.

There are numerous aspects to consider in defining the model in this way. Firstly, the approach was deemed efficient in maintaining large test samples in the sense that no data is lost but rather tested in four different settings with different characteristics. If multi-observation-years had been approach by randomly excluding all but one observation the test sample would have the same size but it would also exclude the information existing in the observations removed. Instead the approach applied has merely moved these observations to another sub-set that has also been tested. Secondly, in separating the sample based on the ESG-components and objective result rather than based on any other criteria allows for more detailed analysis in the sense that the model allows for different effects between the four types of engagements.

This approach removes the ability to test for the engagements as a whole and is not able to fully capture the effect that engagement in a non-specified form may have. Further, one could argue that choosing a panel model and thereby having to reduce the sample size and alter the data set in the ways explained could have been avoided by choosing another methodology, such as using a multivariable OLS regression model for example. Having the data set structured in a way that allows for a panel data model, the benefits and characteristics of such a model as mentioned above outweighs the loss in number of observations by the increased efficiency and detail of the estimates. It has been deemed

more relevant and informative towards the purpose of this study to break the data set into the ESG components as presented.

4.3.2.1 Hypotheses model 2a

The hypotheses that will be tested this using this model are;

(10a) H_0= Engagements do not affect ROA significantly different from zero

H_1 = Engagements have a non-zero effect on ROA

(10b) H_0= Engagements do not have lagged effects on ROA significantly different from zero, in time periods t+1 or t+2

H_1 = Engagements have a non-zero effect on ROA with lagged effects in time periods t+1 or t+2

(11a) H_0= Engagements do not affect Tobin's Q significantly different from zero

H_1 = Engagements have a non-zero effect on Tobin's Q

(11b) H_0= Engagements do not have lagged effects on Tobin's Q significantly different from zero, in time periods t+1 or t+2

H_1 = Engagements have a non-zero effect on Tobin's Q with lagged effects in time periods t+1 or t+2

(12a) H_0= Engagements do not affect Sale per employee significantly different from zero

H_1 = Engagements have a non-zero effect on Sale per employee

(12b) H_0= Engagements do not have lagged effects on Sale per employee significantly different from zero, in time periods t+1 or t+2

H_1 = Engagements have a non-zero effect on Sale per employee with lagged effects in time periods t+1 or t+2

4.3.3 Model 2b - Corporate Financial Improvement

To complement the perspective offered from the model above, some slight adjustments has been made to the model to more closely replicate model 1. This is done by creating binary variables of the dependent variables called "Delta ROA_{t+1}" and "Delta ROA_{t+2}" in the case of ROA but the same for all dependent variables Y. Delta Y_{t+1} takes a value of 1 if the performance measure Y has increased from the year of the engagement to the year after and 0 otherwise. Similarly, Delta Y_{t+2} takes the value 1 if Y is higher year t+2 than at time *t* and 0 otherwise. Defining the variables this way brings two important characteristics when interpreting the results. Firstly, such a probit model does not measure

the absolute effect of a change from 0 to 1 in the X variable of interest but rather measures the probability of the dependent variable taking the value of 1, given the regressors in the model. This means that this model targets the question whether engaging with a firm increases the probability of an improved ROA, Tobin's Q or Sales per employee. Thus, the magnitude of such change is ignored and the importance is whether the engagement positively, negatively or neutrally influences the performance measure, to investigate whether there are any operational improvements stemming from engagements. Secondly, this model addresses the aspect of timing somewhat differently. All variable data has been retrieved from the last of December in the given year. This means that when comparing from year t to year t+1 the model is isolating a time period of exactly one year, as compared to when the tests are done in model 2a. Such isolation can be important and because the engagements can be done throughout the year, with the variable data taken from the end of the same year, there is a difference between practically each engagement in the potential effect it might have up until the first data observation. What this model accomplishes is looking at the event window starting the last day of the year of the engagement and *exactly* one year ahead for the t+1 tests and *exactly* two years ahead for the t+2 model. This is different from the other models and an important distinction to make when interpreting the results.

To more specifically define the model and its characteristics, the reader is directed to section 4.3.1 that discusses Model 1. Model 2b is essentially the same with the subtle difference that the dependent variables are binary rather than ordered 0-3. As mentioned, having an ordered probit model brings up some items that needs attention, such as if the distance between each discrete number is identical or assumed to be non-linear. This is not a concern with a binary probit model which simplifies the interpretation but does not change the structure or usefulness of the model itself.

4.3.3.1 Hypotheses model 2b

Model 2b will test the following hypotheses;

(13) H_0= Engagements do not improve ROA in time periods t+1 or t+2 after an engagement

H_1 = Engagements improve ROA in time periods t+1 or t+2 after an engagement

(14) H_0= Engagements do not improve Tobin's Q in time periods t+1 or t+2 after an engagement

H_1 = Engagements improve Tobin's Q in time periods t+1 or t+2 after an engagement

(15) H_0= Engagements do not improve Sale per employee in time periods t+1 or t+2 after an engagement

H_1 = Engagements improve Sale per employee in time periods t+1 or t+2 after an engagement

4.3.4 Model 3 - Excess return

Presented in section 4.1.5 was a description of how the excess return has been calculated. To make use of this, this model sets out to explore if there are tendencies of non-zero excess returns and whether there are differences between companies' return in years with engagement and the returns of firms not engaged in that year. Stated differently, is there a significant difference in the returns from engaged firms compared to non-engaged ones? Does an engagement have to be successful for this relation to occur or are the effects similar for all engagements? Do the stock market value engagements on any one of the ESG-components more than for the others? Before investigating and comparing engaged companies to the non-engaged, it should be established if the engaged companies have over performed and thereby generated excess returns. It is interesting to measure this not only immediately in the year of the engagement but also one and two years afterwards to see how time plays into the relation. To be able to identify and prove the existence of non-zero returns, a two-tailed one-sample t-test will be conducted. The one sample t-test is a common statistical procedure used to determine whether a sample of observations could have been generated by a process with a specific mean. The one sample t-test estimates the ability to reject the null hypothesis or whether the alternative hypothesis could be accepted. A two-tailed alternative hypothesis will be used in this model and assumes that some difference exists between the true mean (μ) and the comparison value which in this case is zero. The null hypothesis assumes that the difference between the true mean and the comparison value is equal to zero (Stock & Watson, 2012). Thus;

$$H_0: \mu = 0 \qquad \text{and} \qquad H_1: \mu \neq 0$$

As a parametric procedure, the one sample t-test relies on a few easily addressed assumptions;

- That the values are measurable, meaning that they are either continuous or ordinal in their definition which is true for this model as the excess returns are continuous.
- The sample should be representative of the population it is taken from which is deemed to be the case for the samples constructed.
- In efficiently estimating the t-value, the size and normality of the sample is of importance. This will be addressed in more detail below
- That the variance across the samples are intuitively thought to be similar, enabling testing with equal variance.

With the assumptions in mind, the t-statistic can be estimated following four steps (Stock & Watson, 2012);

(1) $$\bar{y} = \frac{(y_1 + y_2 + \cdots y_n)}{n}$$

(2) $$\sigma = \sqrt{\frac{((y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \cdots (y_n - \bar{y})^2)}{n-1}}$$

(3) $$t = \frac{(\bar{y} - 0)}{\frac{\sigma}{\sqrt{n}}}$$

(4) $$p = 2 * \Pr(T > |t|) \; (for\ a\ two - tailed\ comparison)$$

Where equation one estimates the sample mean, $\bar{y}$, by dividing the sum of all observations divided by the number of observations in the sample. The second equation estimates the standard deviation in the sample. That is the square root of the sum of the squared difference between the observed values of y and $\bar{y}$, divided by the number of degrees of freedom, n-1. The t-value in equation three is then calculated by using these two measures and more specifically by dividing the difference between the sample mean and the comparison value (in this case 0) by the standard deviation divided by the number of observations. The t-value is then applied to equation four where it is related to the p-value by taking two times the probability that a random variable within the assumed distribution T is larger than the observed t-value. The calculated value will give the probability (p-value) of observing the test statistic under the null hypothesis. Whether the results provide sufficient evidence to reject the null hypothesis in favor of the alternative hypothesis is determined by looking at the p-value. A lower p-value means a lower probability of obtaining a result like the one defined in the null hypothesis and therefore also means a higher probability of having a random observation deviate from the value in the null. The cutoff value for determining the statistical significance will be at the 10% confidence level, which is to be seen as a less conservative. The same 10% level will also be used for the further tests' being conducted in this model.

The data set has been set-up in the following way. Starting with the full sample of firms that have been engaged at least once in the period between 2005-2016, companies who have been engaged multiple times within the same year has been reduced to solely one observation. Within the individual sub-samples E, S or G, the engagement related to that ESG-component has been prioritized, just like in previous models explained. By sorting the data this way, it is possible to conduct one-sample t-tests with the advantage of discovering isolated effects that stems from the engagements effects on year-on-year and lagged returns and within each sub-sample independently.

The next part of the model addresses the question whether there is a statistically significant difference in the returns from engaged firms compared to non-engaged ones. Along these lines, the data set has been modified in multiple steps to allow for such a comparison. More specifically, for each year two groups have been formed. Based on the full sample of firms that have been engaged at least once during the interval of 2005-2016, let the sub sample x_t contain all firm observations where the firm has been engaged and follows specific criteria in year t and y_t be all other firms remaining in the sample (that has not been engaged in that year). Assigning each firm observation in year t to either group x or y allows for comparison of the average excess returns between these groups and any patterns to this can be identified through a t-test. To estimate a difference between two means (μ_X and μ_Y in two separate samples), a two-sided null hypothesis is formed saying that the means of the two samples differ by a given amount, *d*. Thus;

$$H_0: \bar{x} - \bar{y} = d \qquad \text{and} \qquad H_1: \bar{x} - \bar{y} \neq d$$

This is tested by using the average excess returns in the two groups, $\bar{x}$ and $\bar{y}$. Because these two samples are randomly constructed each year, they are independent random variables. This is an important feature as it combined with the fact that according to the central limit theorem the sample averages are approximately distributed with N(μ, σ^2/n) allows for the t-statistic to be estimated as follows (Stock & Watson, 2012);

$$t = \frac{(\bar{x}-\bar{y})}{SE(\bar{x}-\bar{y})} \qquad \text{where} \qquad SE\ (\bar{x}-\bar{y}) = \sqrt{(\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y})}$$

This method allows for testing in each year separately and thus the effects can both be isolated, but also be lagged to investigate how the effect holds over time and in different periods. If the t-statistic suggests that the null hypothesis does not hold, there is a significant difference between the average excess returns and by using a two-sided approach, this can be both positive and negative. An important aspect in this model is the complementary analysis that can be drawn from the tests. Estimating differences in excess returns depending on engagement criteria is central, but as proposed chapter two, risk reduction is seen as a complementary benefit of ESG-activities for investors. Comparing the standard deviation between the two groups is therefore another area of interest and in doing this it will be possible to deepen the analysis and either further strengthen the positive effect or reduce any

negative effects. By defining the sub samples on different criteria such as only achieved engagements or per ESG-component generates a fuller picture of the relationship between the ESG-engagements and their potential to influence stock prices. In these tests, equal variance is assumed across the samples. This is based on the logical reasoning that all firms are part of the same original population and that a random sub sample-arguably has the same characteristics as the population.

There are some limitations as to what can be established using this method. The first limitation relates to the data itself and how it can be segmented. The more narrowly specified the segments, the fewer observations will meet those requirements and this can limit the usefulness of this model. Both in the case where no observations in a given year matches the requirements but also when the number of observations are low, the statistical accuracy can falter. Following the arguments of the central limit theorem, the quality of the normal approximation relies on the distribution of the observations and if the sample distribution is far from normal, samples with n<30 might not provide accurate estimations (Stock & Watson, 2012). Figure eight is a representation of the full sample and the distribution of excess returns for all firms in the year of engagement. This has been considered when assessing whether a certain segmentation can be accurately tested or not.

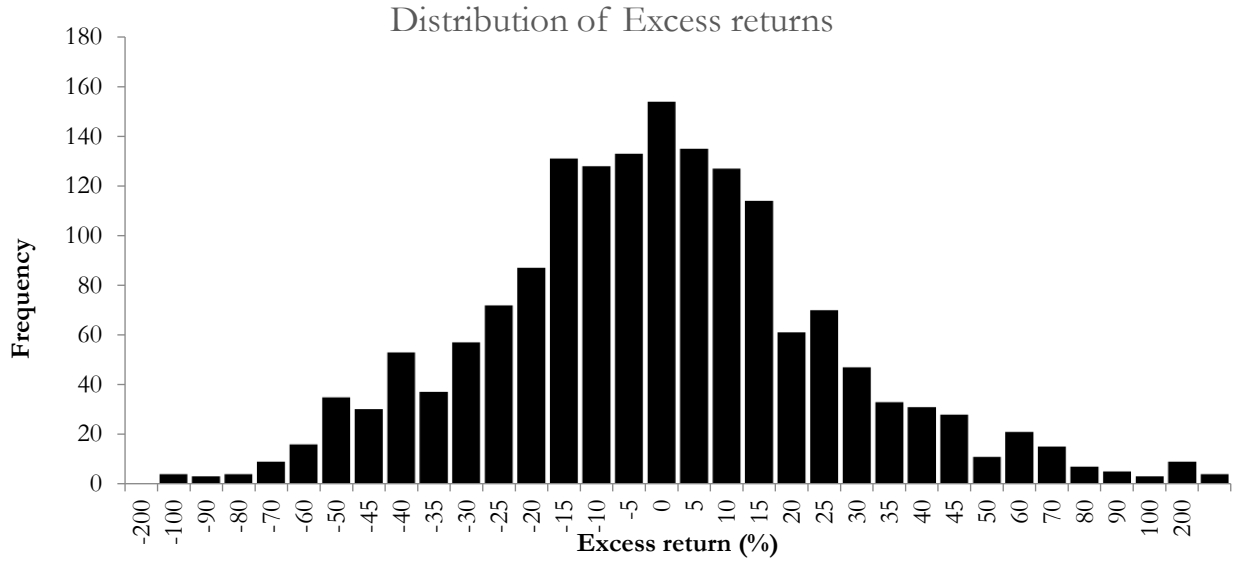


Figure 8: Distribution of excess returns. This histogram shows the distribution of the excess return estimates that has been formed in accordance to section 4.1.5. The frequency represents the number of observations in each bin, and the bins are increasing by increments of five on the horizontal axis. The purpose of the graph is to illustrate the normality of the distribution.

Secondly, as mentioned more in detail above, only firms with available market data for the entire time interval has been included in the tests. Excluding firms that are missing data can create a survivorship bias, meaning that only stable, profitable and "good" firms are incorporated in the analysis and there

through biasing the results. The model has been specified in such a way that limits these effects by not measuring the excess returns compared to the market in general but compared to a benchmark from the same pool of firms. This means that any of the excluded firms could be equally likely to have been present in the group x as they could have been in y, making their exclusion less of an issue. A third area of attention is whether the effects of previous engagements are included in the subsequent year's return estimates. This could be the case if for example a firm has engagements several years in a row and that the potential change in return in year t is not stemming from the engagement in year t but rather in year t-1 or t-2. Not being able to consider autocorrelation, this model could be adjusted by future studies to include such analysis and to see how long lasting the stock price effects of engagements are.

4.3.4.1 Hypotheses Model 3

This model will test the following hypotheses:

(16a) H_0= Engaged firms have not been able to produce returns significantly different from zero

H_1= Engaged firms have been able to produce returns significantly different from zero

(16b) H_0= Engaged firms have not been able to produce returns significantly different from zero in lagged periods t+1 or t+2

H_1= Engaged firms have been able to produce returns significantly different from zero in lagged periods t+1 or t+2

(17a) H_0= The test group of engaged firms has not produced significantly different average returns than the control group of non-engaged firms

H_1= The test group of engaged firms has produced significantly different average returns than the control group of non-engaged firms

(17b) H_0= The test group of engaged firms has not produced significantly different average returns than the control group of non-engaged firms also with lagged periods t+1 or t+2

H_1= The test group of engaged firms has produced significantly different average returns than the control group of non-engaged firms also with lagged periods t+1 or t+2

5. Results

In this chapter the results from the models outlined in the previous chapter will be presented and the interpretation of these will be coupled with brief explanations of the relations discovered. The structure will follow that of the models where each model is presented individually and the results are not analyzed but presented with descriptions of how to interpret the results where necessary. The purpose is to present the findings in an objective and structured way so that the reader can refer back to the text and tables in this section throughout the subsequent chapters as well.

5.1 Model 1 - Objective result

The complete results from model 1 can be seen in table 13-19. In interpreting the table, there are two important aspects to consider. First, the table presents the marginal effect of the variables, thus providing the reader with the estimated change in the probability that the engagement was achieved given that the variable criteria is met. Secondly, the direction of the relationship is potentially of greater practical value than the absolute effects. A negative relation is to be interpreted as negative for the probability of achieving the change related to that particular engagement in relation to the average. A negative marginal effect of X is not indicating that a certain engagement has a negative probability of being achieved, but rather that the probability of it being achieved is X% lower than the average (non-specified) engagement. The model is repeated for four different samples, first all objective engagements were included and additionally sub samples of environment, social and governance engagements were tested individually. The results however, are to be seen and analyzed jointly as they together provide a more detailed picture of the relations.

When testing the effects of the different ESG components, they all provide different results. Environmental engagements proved to have a significantly positive impact on the probability of the objective being achieved at the 5% significance level. Although the case might seem similar for social engagements, the magnitude of the effect is rather weak and the model is not possible to find any significant difference from zero, meaning that social engagements on average does not have a higher or lower probability of being achieved than the sample mean. Contrastingly, engagements on governance topics prove to have a significantly negative effect on the probability of an engagement being achieved. This effect proves to be rather strong and indicates that governance changes potentially includes elements that are harder to communicate to the management, harder to implement or perhaps harder to value in terms of benefits and improvements for the firm. Taken together, these

results indicate that there are notable differences in the ease of achieving change objectives across the ESG-components and so the null-hypotheses for hypothesis 1 can be rejected, as there are differences between the ESG-components.

	Whole sample		
ESG-component	ME	SE	Wald $\chi 2$
Environment	0.1893**	0.083	44.5493***
Social	0.0739	0.078	39.4649***
Governance	-0.2140**	0.073	46.1708***
n	**1341**		

Table 12: Model 1: ESG-components. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

Building on this, testing whether certain topics are easier or harder to change than others can contribute to an understanding of what sorts of changes are most commonly implemented by the target firm. Each topic has been tested individually in the entire sample but also in each relevant sub sample in order for the results to be confirmed and further highlighted. On an aggregated level there are four topics that prove to differ significantly in their probability of being achieved. Corporate strategy proved to be positively related to the probability of succeeding with a given change, potentially indicating that in these types of issues there is little discrepancy between what management and investors (Schroders) believe to be in the best interest of the firm. This topic is related to governance engagements and the result, although slightly contradicting to the result of governance engagements being harder to achieve, strengthens the impression that strategic changes might be an area that investors can influence. Business integrity proves significantly positive and this is also a topic related to the governance component of ESG, implying similar implications as for Corporate strategy. However, turning to the topic of Shareholder rights, this relationship is reversed and we see a significantly negative impact on the probability of success. Also being a topic within governance, the differences between these topics can help understanding what distinguishing characteristics might drive the overall negative relation between governance engagements. Having both positive and negative effects of the magnitudes that are seen in this test is highly interesting as it signals an achievability of governance engagements but that there are some factors/topics/circumstances that hinders this, areas that can be highly helpful in driving change if correctly identified. The last topic that proved to be significant in the model is Pollution, creating some of the positive impact that environmental engagements have on the probability of success. This positive effect holds both in the aggregated sample but also when environmental engagements are tested separately. This could be seen

in relation to the topics of Climate change and Water that did not show similar effects. Although such conclusions should be inferred with caution, there might be attributes to engagements on the topic Pollution that are absent in the other engagements.

These results provide some support for the alternative hypothesis (H_1) in hypothesis 2 as the probability of success show tendencies to vary depending on the topic. The insignificant marginal effects of the remaining topics indicate one or two things. First the objective results of these engagements can be largely centered around the sample mean and thus the model is incapable of identifying any significant variations, simply because there are none. Or the objective results could be largely spread out with high success rates and also very low, ultimately averaging out the total effect with almost random outcomes. By looking at the descriptive statistics from the social sub sample one can see that social engagements have a mean probability of success of 45% (translates to 1,77 with the numbering applied in model one) and large variance (a standard error of 1,3), meaning that there are large variations across the sample, potentially removing the effects on average, and suggesting that the later of the two scenarios might prevalent. Thus, the social sub sample is not able to reject the null-hypothesis from hypothesis 2.

	Whole sample			**Environment**			**Social**			**Governance**		
Topics	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2
Water	0.2854	0.199	41.5204 ***	0.1652	0.202	33.7998 ***	-	-	-	-	-	-
Climate Change	0.0854	0.115	39.6014 ***	-0.1197	0.138	33.8027 ***	-	-	-	-	-	-
Pollution	0.4735 ***	0.141	52.4846 ***	0.4136 ***	0.151	39.5274 ***	-	-	-	-	-	-
Human Capital Management	0.0238	0.114	39.0828 ***	-	-	-	-0.0636	0.128	21.1027 ***	-	-	-
Health & Safety	0.0564	0.148	39.1746 ***	-	-	-	0.0311	0.157	20.8824 ***	-	-	-
Customers	0.1109	0.172	39.4169 ***	-	-	-	0.0608	0.176	21.2014 ***	-	-	-
Supply Chain Management	0.0040	0.177	39.0910 ***	-	-	-	-0.0669	0.180	20.8106 ***	-	-	-
Business Integrity	0.2382 **	0.115	44.1201 ***	-	-	-	-	-	-	0.4680 ***	0.132	23.2359 ***
Shareholder Rights	-0.9970 ***	0.133	92.2312 ***	-	-	-	-	-	-	-0.9555 ***	0.148	49.6111 ***
Corporate Strategy	0.3966 ***	0.108	52.0612 ***	-	-	-	-	-	-	0.6737 **	0.125	37.6156 ***
n	**1341**			**348**			**431**			**563**		

Table 13: Model 1: Engagements on different topics. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

When investigating whether engagements are more or less successful depending on how the target is approached, the model suggests that a collaborative engagement is significantly less likely to succeed. This is driven by the fact that the governance engagements show a large negative effect. Also, no objective environmental engagements have been conducted in collaboration with other investors and social engagements do not show any significant deviation from the average success rate. The fact that most engagements done collaboratively are seen in governance where 115 collaborative engagements have been initiated with only twelve times under the social component, indicates that the negative relation might be true for the other subcomponents of ESG as well but the small sample size prevents the model from finding strong relations. Whether engagements are becoming more and more efficient is an important aspect to consider when assessing how likely an engagement is to succeed. To address this question, a variable containing the number of engagements for each firm has been tested against the success variable. It turns out significantly positive, though with a rather small magnitude. Nonetheless, this potentially shows the relational aspect between the parties and the value of working together with change more than once. Breaking it into the ESG components, governance is the only one out of the three that proves a significantly positive relation to the probability of success. These results indicate that the null-hypothesis in hypothesis 3 can be rejected, as there are differences in the probability of success depending on the collaborative and repetitive nature of engagements.

	Whole sample			**Environment**			**Social**			**Governance**		
Engagement approach	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2
Collaborative	-0.9601 ***	0.123	98.3910 ***	-	-	-	-0.1126	0.373	20.7498 ***	-1.0302 ***	0.142	59.9508 ***
Number of engagements	0.0312 ***	0.007	58.7780 ***	-0.0173	0.035	33.0668 ***	0.0133	0.025	20.8979 ***	0.0453*	0.023	55.3476 ***
n	**1341**			**348**			**431**			**563**		

Table 14: Model 1: Collaborative and number of engagements. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

When instead conducting the tests based on geographic regions, the results for the ESG-components shifts somewhat. Governance, showing strong results in various topics and in the engagement approach, proves to be less sensitive to geographical variation. There is a significantly positive effect from engaging in firms in the EMEA region as the probability of success by seems to increase with about 50% if the firm is active in this market. This result holds for all sub-components although the effect is even stronger for environmental and social engagements where governance engagements

have lower impact than the other two. Environment proves to be more sensitive to geographical variation as, apart from the significantly positive effect seen in the EMEA-region, these engagements show a significant decrease in the probability of success when firms are engaged in Asia and North America, with similar magnitude as for the EMEA region. This is the case for social engagements as well, although they also experienced a distinct negative effect from being in the Pacific region. Similar to the case when looking at the effect of collaborative engagements, the lack of clear results for the South America region could be explained by the low number of such engagements (n=8). Taken together there seems to be regions that are more reluctant to implement the suggested changes and so the null hypothesis 4 can to a large extent be rejected.

	Whole sample			**Environment**			**Social**			**Governance**		
Region	ME	SE	Waldχ2	ME	SE	Waldχ2	ME	SE	Waldχ2	ME	SE	Waldχ2
Asia	-0.5490 ***	0.161	49.1747 ***	-0.5117 **	0.260	37.2030 ***	-0.9039 ***	0.318	27.8947 ***	-0.4330	0.273	13.2145 *
EMEA	0.5493 ***	0.092	60.6023 ***	0.6577 ***	0.158	51.1206 ***	0.8528 ***	0.150	47.5765 ***	0.2820*	0.165	13.3407 *
N. America	-0.4325 ***	0.125	49.7781 ***	-0.5628 **	0.236	38.1348 ***	-0.6047 ***	0.184	29.7839 ***	-0.1665	0.243	11.3201 *
S. America	-0.2128	0.473	39.1393 ***	-0.8502	0.915	33.4819 ***	-0.7337	0.735	21.7157 ***	0.7899	0.804	11.8809 *
Pacific	- 0.2738 **	0.139	42.1284 ***	-0.3643	0.230	35.9445 ***	-0.5139 **	0.257	24.4253 ***	-0.1564	0.230	12.2487 *
n		**1341**			**348**			**431**			**563**	

Table 15: Model 1: Engagements on different geographical areas. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

The model was also constructed so that different sectors could be investigated and provide insights into what specific characteristics might be favorable when opting to successfully interact with target firms. These tests work to further strengthen the idea that different settings and external characteristics creates variations in the willingness of firms to incorporate and initiate changes in their ESG-activities. The model suggests that engagements in the consumer staples, materials, telecom services and utilities sectors show a strong positive tendency to achieve the proposed change. The magnitude differs between these as well, with the telecom sector showing the strongest positive effect of about 90% increased probability of being achieved where materials has an estimated 21% increased probability of success, and the remaining two in between this interval. Interestingly, these effects vary across the ESG-components as well. Starting with the most prevalent effect, the positive effect for telecom services is driven solely by engagements on environment and social issues, where the governance component suggests little to no deviation from the mean probability of success. Utilities on the other

hand prove to be beneficial across the segments as all three components prove a strong increase in the likelihood of success related to those engagements. Contrastingly, measured effects for the consumer staples and Materials industries seem to come exclusively from governance engagements.

On the opposite side of the spectrum the model suggests that financials and Information Technology (IT) firms have a strong reluctance to meet the suggested change. The significantly negative effect on the probability of success for financial firms is driven by environmental engagements and although both social and governance engagements seem to experience similar resistance, their effects are not of the same magnitude. Environmental engagements are furthermore contributing largely to the heavy decrease in the probability of success in changes for IT firms. This is also strengthened by a strong negative influence from governance engagements. Aggregating the results, when looking at the whole sample the model clearly suggests that there are variations across sectors as to how successful previous engagements have been. It is therefore an area that is worthwhile exploring further as identification of key characteristics that might be causing these deviations can greatly enhance the ability to encourage appropriate change initiatives in a more sector manner. As for environmental engagements, having both positive and negative effects for engagements suggests that these types of projects are more suitable in certain sectors than others, potentially varying due to aspects such as value chains, customer preferences and sector priorities. The social engagements, interestingly, only prove to be positively or neutrally affected by various sectors, indicating that these changes are either easier and less cumbersome to process, or considered closer to the core business to a larger extent than the other two ESG-components. This could be the case as governance is seen having both positive and negative changes to the probability of success in a given change depending on the sector of the firm. Taken together, these tests provide rigorous proof to reject the null hypothesis 5 as the sector has large impact on the probability of success and thus creates variation across businesses.

	Whole sample			Environment			Social			Governance		
Sector	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2
Consumer Discretionary	-0.1556	0.113	42.2903 ***	0.1703	0.241	33.1906 ***	-0.1971	0.180	22.4809 ***	-0.2211	0.182	14.1567 *
Consumer Staples	0.3309 ***	0.119	43.8939 ***	0.1763	0.252	33.1853 ***	0.2315	0.181	21.3950 ***	0.5043 **	0.202	17.0475 **
Energy	-0.2240	0.140	41.1505 ***	-0.0513	0.239	33.0518 ***	-0.1933	0.253	21.3280 ***	-0.3718	0.231	13.3361 *
Financials	-0.4270 ***	0.115	49.2512 ***	-0.8896 ***	0.213	47.7332 ***	-0.2892	0.206	21.3722 ***	-0.2611	0.184	12.1914 *
Health Care	-0.2709	0.178	41.4769 ***	-0.0464	0.438	33.0437 ***	-0.3042	0.313	21.4198 ***	-0.3002	0.257	12.5124 *
Industrials	-0.0001	0.116	39.0909 ***	0.3109	0.246	33.8920 ***	0.1860	0.195	21.5749 ***	-0.2421	0.180	12.1878 *
Information Technologies	-0.7885 ***	0.178	54.9629 ***	-1.4354 ***	0.365	46.6177 ***	-0.3643	0.282	21.8781 ***	-0.8106 ***	0.294	17.6564 **
Materials	0.2078 **	0.098	44.2118 ***	0.0467	0.157	33.6221 ***	0.0603	0.185	20.8535 ***	0.4315 **	0.169	17.0836 **
Telecom Services	0.9209 ***	0.238	55.1200 ***	1.1718 ***	0.370	42.6541 ***	1.3355 ***	0.518	21.5803 ***	0.4806	0.379	18.8973 **
Utilities	0.8231 ***	0.171	61.7152 ***	0.6143 **	0.258	37.8725 ***	0.8183 **	0.405	24.8596 ***	0.9945 ***	0.285	21.1957 **
n	**1341**			**348**			**431**			**563**		

Table 16: Model 1: Engagements on different sectors. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

The above scenario, where the effects differ largely across the ESG-components is almost vanished when incorporating the three performance measures in the model. ROA turns out to be significantly positive across all components and indicates that healthy firms generating a high return in relation to their assets are more inclined to pursue ESG-implementations within their organizations. This effect is not visible for Tobin's Q that show no significant effect for any of the components. Having a clear relation similar to that of ROA would suggest that not only does a good ability to generate healthy profits encourage the firm to take on ESG-initiatives but also that a high market valuation would do so. Seeing that this is not the case could instead suggest that firms operate based on their own conceptions rather than taking that of the public market into consideration when initiating change projects within ESG. The model also proves that less efficient firms have experienced a slightly higher likelihood of achieving an objective. Sales per employee has negative estimates throughout the tests, yet all but one of the four samples proved to be significantly lower than zero. For environmental engagements, higher sales per employee reduced the probability of achieving the objective where for the other ESG-components and on the aggregate level, this effect is not distinguishable from zero. These results show that for ROA, the null hypothesis in hypothesis 6 can be rejected across the board.

For Tobin's Q however, the null hypothesis in hypothesis 7 cannot be rejected as no significant difference from zero could be proven. This result applies rather strongly to the Sales per employee as well, where the null hypothesis could only be rejected at the 10% level for the environmental sub-sample but the model failed to reject the null of hypothesis 8 in the other tests.

	Whole sample			**Environment**			**Social**			**Governance**		
Performance measures	ME	SE	Wald$\chi 2$	ME	SE	Wald$\chi 2$	ME	SE	Wald$\chi 2$	ME	SE	Wald$\chi 2$
ROA	2.3380 ***	0.416	39.7255 ***	2.5675 ***	0.757	35.6763 ***	1.9419 **	0.802	25.2273 ***	2.1454 ***	0.646	21.5702 ***
Tobin's Q	-0.0282	0.038	41.5234 ***	0.0878	0.088	33.0627 ***	-0.0397	0.066	21.0670 ***	-0.0604	0.054	11.9844 *
Sales per Employee	-0.0408	0.028	34.8881 ***	-0.0997 *	0.052	39.5287 ***	-0.0546	0.050	18.1577 **	-0.0187	0.044	13.6567 *
n	**1341**			**348**			**431**			**563**		

Table 17: Model 1: Engagements on different performance measures. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

Although they are not of primary focus in this paper, the firm specific characteristics have also been tested with the model as separate variables of interest in addition to being used as control variables in the models. Contrasting the results of Dimson et al. (2015) the model suggests that size is not favorable for achieving an ESG-initiative within the average firm. The variable proves to be significantly negative for the whole sample and also for the sub-components of environment and social but it becomes insignificant for governance engagements. This could imply that environmental and social changes to a larger extent affects operational aspects of the firm and that larger firms have more distinguished processes and thus lower ability to change these. The M/B ratio shows some significance in the social component and with a negative relation, the result suggests that when market expectations on the firm are higher, the less likely a firm is to go achieve a social objective. Possibly the most influential firm characteristic is sales growth as it proves to positively influence the probability of success across all the samples. Having growing sales indicates a healthy and expanding business amongst other things and initiating ESG-changes might be relatively easy for firms as their operations are expanding and thus already are under some development, easing the inclusion of ESG-initiatives. The model suggests that R&D and CapEx spending are somewhat affecting the probability of succeeding with an engagement. Governance engagements are positively affected by high levels of R&D in relation to total assets whereas the success of environmental engagements is negatively affected by higher capital expenditures. Contrasting the two, it can be argued that governance engagements require lesser capital investments whereas environmental engagements are more capital intensive in their nature. This

concludes that also the null hypothesis of hypothesis 9 can be rejected as different firm specific characteristics does affect the probability of succeeding with an engagement, indicating that some type of firms might be more easily targeted than others.

	Whole sample			Environment			Social			Governance		
Firm characteristics	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2	ME	SE	Wald χ2
LnAssets	-0.0700 ***	0.016	39.0950 ***	-0.0840 ***	0.030	33.0512 ***	-0.1145 ***	0.031	20.7482 ***	-0.0294	0.026	12.7813 *
D/E	0.0000	-	39.0950 ***	-0.0001	-	33.0512 ***	0.0001	0.000	20.7482 ***	0.0000	0.000	12.7813 *
Cash/Assets	-0.5780	0.444	39.0950 ***	-0.4854	0.849	33.0512 ***	-0.6756	0.719	20.7482 ***	-0.4080	0.738	12.7813 *
M/B	-0.0041	0.003	39.0950 ***	-0.0032	0.006	33.0512 ***	-0.0097 *	0.006	20.7482 ***	-0.0015	0.004	12.7813 *
Sales Growth	0.4363 ***	0.101	39.0950 ***	0.8442 ***	0.215	33.0512 ***	0.5404 **	0.225	20.7482 ***	0.2962 **	0.135	12.7813 *
R&D/Assets	2.0712	1.324	39.0950 ***	0.9432	2.881	33.0512 ***	1.4616	2.369	20.7482 ***	3.4270 *	1.933	12.7813 *
CapEX/assets	-0.5214	0.856	39.0950 ***	-2.6047 **	1.521	33.0512 ***	-1.3518	1.505	20.7482 ***	0.9297	1.426	12.7813 *
n	**1341**			**348**			**431**			**563**		

Table 18: Model 1: Engagements on different firm characteristics. This table presents the results from Model 1. The coefficients are the marginal effects estimated for each variable, indicating that it is by that amount the probability of achieving an objective increases or decreases with. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All models have been used with the control variables presented in section 4.2.

5.2 Model 2 - Corporate Financial Performance

Taken together, in the tests conducted, this model does not seem to present many relationships that indicates significant effects from the engagements done. In this section, table 20 and 21 presenting the results from each test serves as a complement to the descriptions provided in the text below. Starting with model 2a, measuring the effects of ESG-engagements on ROA, Tobin's Q and Sales per employee with two year lags there are no significant relations presented from the model. Interpreting the environment subsample, the immediate effect on ROA has a negative impact and although this is not significantly different from zero, it indicates that within the same year of an engagement the returns of an environmental investment might not be easy to realize. This relation changes to a positive but still insignificant result when including both one and two years lag. Looking at the coefficient estimates for environmental engagements and the effect on ROA, the interpretation is that having an engagement in year t on average reduced the ROA with 0,61%. In year t+1 the same engagement had a positive influence on ROA and on average resulted in a 0,08% increase, with year t+2 having a larger increase of 1,4%. Had these coefficient estimates been significantly different from zero, these results would have enabled further analysis on why the effects change over time and what implications they might have. However, with the standard errors being rather large, the model is unable to dictate whether there is any effect at all or whether it is positive or negative due to the fact that the variation across the observations creates wide confidence intervals. Looking more thoroughly at the standard

errors, one should note that these are lower for all three ESG-components in the year t+1. With lower standard errors, there is less variation in the observations and even though they do not deviate significantly from zero in the effect, the effects are more consistent in year t+1 than the other two alternatives. The tests on the social and governance sub-samples also proved insignificant with some of the coefficient estimates from the governance test being negative and the social estimates all positive. The results from these two sub-samples should be interpreted in a similar fashion to the environment sub-sample. It is also recognized that the social engagements have lower standard errors and that the governance and environment samples both experienced larger variance. Whether or not a specific engagement has been an objective and achieved does not change the results significantly. The effect on ROA still remains around zero for the engagements that have been achieved and thus succeeding with a given change does not necessarily have greater impact on ROA than any other sort of engagement. Hypotheses 10a and 10b are constructed with the null hypothesis being that there is no change in ROA, given an engagement in year t, in periods t, t+1 or t+2 and the tests are not able to reject that null for any of the sub-samples. The inability to reject the null for all tests could indicate that engagements on average do not increase the return for a firm given the same assets or that the changes are seen on a much longer time horizon, not addressed with this sample.

Changing the dependent variable to Tobin's Q the results are not greatly changed in terms of significance and the interpretations are similar. The social sub-sample did result in a significant coefficient for the year t+2. This coefficient, 0,0705, is significant at the 10% confidence level and means that when a social engagement was done, firms in the sample on average experienced an increase in Tobin's Q of about 7% two years afterwards. Experiencing increased Tobin's Q values, firms enjoy a higher market value of their assets, which indicates that social engagements contain some characteristic that is valuable to the market. All estimates are positive in the social sub-sample for Tobin's Q and the governance sub-sample shows similar traits, with solely positive coefficients. The environment sub-sample differs from the other two in that the two lagged years have negative coefficients. Worth noting is the fact that for the social and governance sub-samples the coefficients grow in magnitude from year to year with the largest coefficients in year t+2 but that this is coupled with an increase in the standard errors as well. The standard errors are lowest in year t for all ESG-components in this case, proving lower variation across the observations in this year. Also in these tests, the governance and environment sub-samples indicate greater variance in the estimates than is the case for the social sample. Achieved engagements show somewhat different patterns. The estimates remain insignificantly different from zero, but in period t and t+1 the estimated effect of an

achieved engagement is practically identical, indicating that there the time effect is not creating additional differences in the effects. In summary, these tests find little ability to reject the null hypotheses of hypotheses 11a and 11b as practically all estimates turned out insignificant which affirms the weak relation between engagements and CFP in this model.

The results indicate that the Sales per employee measures are not greatly affected by the engagements done. There is one significant result, in the social sub-sample, indicating that social engagements on average resulted in a significant reduction in Sales per employees one year after the engagement. This reduction, of about (-)10,5%, is significant at the 5% confidence level. Generally, the test results indicate a slightly more negative effect on Sales per employees than for the other dependent variables tested yet the majority of the results are not significantly different from zero. The magnitude of the coefficients as well as for the standard errors are relatively large and indicates a great variation in the effects that engagements of all sorts have on the firms and their sales efficiency. Also for this variable, the successful achievement of an engagement does not improve the results as all estimates in the achieved sub-sample are insignificant. Noting that the success is insignificant in improving employee productivity communicates that an engagement that is addressed but where no change is registered is not underperforming one that is achieved. Notable is that the sample size for the tests on this variable is smaller due to less data available from the resources, leading to fewer firm year observations however the samples are considered to be sufficiently large. Nonetheless, the inability to reject the null hypotheses remains true for most part of this definition of the model, failing to reject hypotheses 12a and 12b.

The ability to achieve a given engagement do not prove to be of significant value in this setting either. The estimates are insignificant for both lagged periods, both being negative. Considering this in combination with the previous tests on ROA it seems that there is no benefit from achieving an objective engagement compared the other alternatives, signaling that the effort and reward relation might be asymmetric. With these results in mind, there are little support for the alternative hypothesis of hypothesis 13. Instead, the majority of the tests are unable to reject the null of no change in ROA. Conducting the same test on changes in Tobin's Q, the results prove to be different from each of the ESG-components. For the environment component, the two coefficient estimates are negative, with the estimate of two years lag time proving to be statistically significant. This effect, significant at the 1% level indicates that an environment engagement reduces the probability of experiencing an increase in Tobin's Q with almost 20% two years after the engagement year. As for the social sub-sample there

are no significant relations and with a negative coefficient in the year after the engagement and a positive estimate two years after, the results further indicate that the effects are moving around neutral implications. Contrastingly, the governance sub-sample show positive coefficients in both the lagged years, though these are not significantly different from zero.

The importance of achieving the engagement objective continues to be of low importance as the coefficients are insignificant and move from positive to negative over time. It is evident that the majority of the tests are unable to reject the null hypothesis of hypothesis 14 as only one out of six estimates was significantly different from zero. Lastly, the results from testing the Sales per employee variable are also varying across the ESG-components. As for governance engagements, the coefficients again turn out positive but insignificant whereas this direction is reversed for the social sub-sample where both estimates are negative and statistically insignificant. The environmental engagements move from negatively affecting the probability of an increased amount of Sales per employee in the first year, to a positive effect two years after the engagement. These estimates are however insignificant. Insignificant are also the last estimates testing the achieved sub-sample. This concludes that the insignificant effect that achieving the engagement has, holds across all performance variables and for the two different model definitions. With this, also for hypothesis 15, the tests are unable to reject the null hypothesis. In general, model 2b implies greater variance in the dependent variables compared to model 2a, with a tendency of larger standard errors and thus large confidence intervals.

	Environment				Social				Governance			
	Coeff.	SE	R2	SER	Coeff.	SE	R2	SER	Coeff.	SE	R2	SER
ROA	-0.0061	0.008	0.717	0.002	0.0012	0.007	0.717	0.002	-0.0018	0.007	0.717	0.002
ROA t+1	0.0008	0.007	0.761	0.002	0.0014	0.006	0.761	0.002	0.0062	0.007	0.761	0.002
ROA t+2	0.0140	0.009	0.721	0.002	0.0010	0.008	0.718	0.002	-0.0135	0.009	0.720	0.002
n **(t,t+1)**; (t+2)						**1048** ; *720*						
Number of cross sections						548						
Time series length						11						
Tobin's Q	0.0031	0.042	0.940	0.072	0.0421	0.037	0.939	0.073	0.0156	0.040	0.940	0.072
Tobin's Q t+1	-0.0081	0.043	0.959	0.074	0.0601	0.038	0.959	0.074	0.0188	0.041	0.959	0.074
Tobin's Q t+2	-0.0222	0.045	0.934	0.065	0.0705*	0.040	0.935	0.064	0.0236	0.043	0.934	0.065
n **(t,t+1)**; (t+2)						**1130** ; *778*						
Number of cross sections						603						
Time series length						11						
Ln (Sales/Employee)	-0.0197	0.023	0.993	0.016	-0.0002	0.021	0.992	0.016	-0.0001	0.022	0.993	0.016
Ln (Sales/Employee) t+1	-0.0431	0.048	0.970	0.071	-0.1046**	0.053	0.952	0.104	0.0592	0.045	0.970	0.071
Ln (Sales/Employee) t+2	-0.1142	0.070	0.950	0.115	0.0117	0.063	0.946	0.115	0.0924	0.062	0.95	0.115
n **(t,t+1)**; (t+2)						**739** ; *579*						
Number of cross sections		359				330				358		
Time series length						10						

	Achieved			
	Coeff.	SE	R2	SER
ROA	0.0077	0.007	0.722	0.002
ROA t+1	0.0007	0.006	0.760	0.002
ROA t+2	-0.0111	0.008	0.719	0.002
n **(t,t+1)**; (t+2)		**1048** ; *720*		
Number of cross sections		t=603, t+1=548, t+2=349		
Time series length		t,t+1=11, t+2=10		
Tobin's Q	0.0321	0.039	0.940	0.072
Tobin's Q t+1	0.0309	0.04	0.960	0.072
Tobin's Q t+2	-0.0504	0.041	0.939	0.060
n **(t,t+1)**; (t+2)		**1130** ; *778*		
Number of cross sections		t,t+1=603, t+2=384		
Time series length		t,t+1=11, t+2=10		
Ln (Sales/Employee)	0.0311	0.021	0.993	0.016
Ln (Sales/Employee) t+1	-0.0105	0.045	0.971	0.072
Ln (Sales/Employee) t+2	-0.0659	0.065	0.951	0.117
n **(t,t+1)**; (t+2)		**739** ; *579*		
Number of cross sections		t=369, t+1=361, t+2=280		
Time series length		t,t+1=10, t+2=9		

Table 19: Model 2a: Corporate financial performance –ESG & Achieved. This table presents the results from Model 2a. The model is a panel data model and two way fixed effects have been used. The number of observations differ due to sample optimization where as many observations as possible has been included. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All tests include the control variables presented in section 4.2, except for the tests on Tobin's Q where the M/B ratio has been left out.

	Environment			Social			Governance		
	Coeff.	SE	Wald χ2	Coeff.	SE	Wald χ2	Coeff.	SE	Wald χ2
Delta ROA t+1	0.0223	0.051	40.9329 ***	0.0890 *	0.053	43.5922 ***	-0.0928*	0.048	44.5106 ***
Delta ROA t+2	-0.0513	0.060	71.6849 ***	0.0175	0.066	71.1160 ***	0.0368	0.060	71.3948 ***
n **(t,t+1)** *(t+2)*	**2874** ; 1992								
Delta Tobin's q t+1	-0.0613	0.050	109.6139 ***	-0.0104	0.052	108.2439 ***	0.0619	0.047	109.8658 ***
Delta Tobin's q t+2	-0.1928***	0.058	142.2747 ***	0.0671	0.065	133.8141 ***	0.1362	0.058	137.3691 ***
n **(t,t+1)** *(t+2)*	**3054** ; *2128*								
Delta Ln (Sales/Employee) t+1	-0.0486	0.059	28.3192 ***	-0.0021	0.066	27.8874 ***	0.0494	0.059	28.6410 ***
Delta Ln (Sales/Employee) t+2	0.0276	0.064	15.4662 **	-0.0660	0.072	16.1495 ***	0.0270	0.067	15.4322 **
n **(t,t+1)** *(t+2)*	**2035** ; *1671*								

	Achieved		
	Coeff.	SE	Wald χ2
Delta ROA t+1	-0.0174	0.091	13.9643*
Delta ROA t+2	0.0089	0.111	27.1613***
n **(t,t+1)** *(t+2)*	**1047** ; *719*		
Delta Tobin's q t+1	-0.0624	0.089	26.4406***
Delta Tobin's q t+2	0.0237	0.108	39.1444***
n **(t,t+1)** *(t+2)*	**1129** ; *777*		
Delta Ln (Sales/Employee) t+1	-0.0625	0.110	7.806
Delta Ln (Sales/Employee) t+2	0.0885	0.121	6.1854
n **(t,t+1)** *(t+2)*	**738** ; *580*		

Table 20: Model 2b: Corporate financial improvement - ESG & Achieved. This table presents the results from Model 2b. The probit model estimates the changes in probability that the dependent variable takes on a value of 1, in this case signaling an improvement in the underlying variable from the engagement done in year t. The number of observations differ due to sample optimization where as many observations as possible has been included. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. All tests include the control variables presented in section 4.2, except for the tests on Tobin's Q where the M/B ratio has been left out.

5.3 Model 3 - Excess return

The results addressing the hypotheses that were introduced in section 4.3.4.1 will be presented in this section. Model three started off by testing if companies engaged on the three sub-components E, S and G in a given year had generated any excess return that is statistically significant on a 10% level. Testing the sub-samples one by one will indicate if there exist any isolated effects from the different types of engagements. The results will be of further interest when analyzed alongside the results from the comparison between the engaged and non-engaged companies.

The companies engaged on a governance-topic had the highest number of significant results with approximately two thirds having average returns significantly different from zero and with both positive and negative average returns. Socially engaged companies had most negative t-statistics across all periods, meaning their average returns were below zero, yet they had the lowest number of significant results, making them indistinguishable from zero. From a maximum of twelve tests on the periods between 2005-2016, conducted on excess returns the same year a company was engaged, a little less than half of the periods proved to be statistically significant in all three sub-samples, meaning the null hypothesis in hypothesis 16a cannot be rejected in full but for about half of the individual tests. The number of observations in every period differ between sub-samples, averaging around 30 observations which is acceptable according to the central limit theorem. An evident pattern emerges when looking at lagged returns one and two years after an engagement. Across eleven periods 2005-2015, lagged one year, on all the three sub-samples, there are more significant results than insignificant, with the majority being negative. Environmental engagements proved to be significantly different from zero in the most periods. Estimating the effects two years after an engagement, testing on engagements that occurred between 2005-2014, gives a total of 30 independent t-tests, ten in each sub-sample. Half of these tests proved to be significantly different from zero and all periods except for one was negative. Thus the returns show tendencies of becoming more negative with time. Consequently, the null hypothesis in hypothesis 16b, that there exist no returns significantly different from zero would be rejected in the majority of periods in favor for the alternative hypothesis[11].

[11] Please refer to the appendix 9.3 for a graphical presentation of the excess returns on Environmental, Social and Governance

Excess return: Environmental

Engagement Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Year-on-Year												
Mean	4.16	20.56	47.68	-84.72	18.24	-77.11	-46.38	-1.28	-10.41	-15.61	-14.82	20.60
Std.dev	24.35	32.07	38.34	13.59	49.82	27.64	22.33	2.75	44.27	18.96	28.23	35.71
n	20	32	30	30	33	24	42	31	28	22	37	32
t-stat	0.76	36.26 ***	0.68	-34.15 ***	21.03 **	-13.66	-13.46	-0.25	-12.44	-3.86 ***	-3.19 ***	0.32
t+1												
Mean	1.29	13.12	-53.00	13.26	-97.17	-61.21	-20.69	-15.91	-11.99	-17.07	-63.17	
Std.dev	26.41	35.60	16.46	27.51	33.09	22.62	23.59	34.46	30.88	33.12	2.03	
n	20	32	31	30	34	24	42	31	29	22	36	
t-stat	21.79 **	20.85 **	-17.92*	26.39 **	-17.12 *	-13.26	-0.56	-25.71 **	-20.91 **	-24.17 **	-18.64 *	
t+2												
Mean	10.62	-69.62	92.01	-86.16	-11.06	44.16	-79.91	-11.89	-12.61	27.04		
Std.dev	37.97	14.84	56.77	19.50	22.57	31.69	41.39	8.72	29.19	6.15		
n	20	32	31	30	34	24	42	31	29	22		
t-stat	12.50	-26.54 **	0.90	-24.19 **	-28.56 ***	0.68	-12.51	-3.53 ***	-23.26 **	20.61 **		

Table 21: Model 3: Excess returns: Environmental This table presents the results from Model 3. The tests are one sample t-tests. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%.

Excess return: Social

Engagement Year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Year-on-Year												
Mean	-32.96	86.36	-12.92	-43.50	-10.68	-53.15	-64.01	16.23	-14.85	-15.79	-84.39	-53.61
Std.dev	21.33	17.53	30.65	19.57	43.67	20.69	22.31	35.80	28.83	18.70	3.53	28.51
n	22	15	16	13	21	22	15	23	21	31	101	86
t-stat	-0.72	19.07 *	-0.17	-0.80	-11.20	-12.04	-11.11	0.21	-23.60 **	-0.47	-23.9 9 **	-17.44 *
t+1												
Mean	85.96	0.98	-54.07	25.35	-83.16	16.42	-17.27	-2.20	-12.90	-16.29	13.24	
Std.dev	25.77	25.68	13.64	67.70	0.33	20.98	27.96	47.69	20.10	26.83	55.19	
n	22	15	16	13	22	22	15	23	21	31	102	
t-stat	15.64	0.14	-15.85	13.50	-11.86	0.36	-23.92 **	-22.13 **	-29.40 ***	-33.80 ***	24.22 **	
t+2												
Mean	-90.47	-2.23	-1.16	-74.33	-14.18	45.46	-28.02	-12.49	-12.69	53.78		
Std.dev	26.96	10.07	37.66	21.82	20.72	22.31	5.74	30.27	22.14	37.14		
n	22	15	16	13	22	22	15	23	21	31		
t-stat	-15.73	-0.85	-12.30	-12.28	-32.10 ***	0.96	-18.91 *	-19.78 *	-26.25 **	0.81		

Table 22: Model 3: Excess returns: Social. This table presents the results from Model 3. The tests are one sample t-tests. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%.

Excess return: Governance												
Engagement Year	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Year-on-Year												
Mean	75.32	10.84	12.16	51.10	1.42	-38.21	-13.41	13.03	-16.85	-13.62	0.00	-35.51
Std.dev	17.75	23.60	35.31	16.28	32.07	18.79	22.36	0.00	39.18	22.09	27.21	3.09
n	20	19	13	8	18	23	30	23	31	105	224	411
t-stat	18.97 *	20.03 *	12.42	0.89	18.74 *	-0.98	-32.84 ***	25.45 **	-23.94 **	-63.19 ***	0.13	-23.32 **
t+1												
Mean	11.19	-72.29	-11.41	39.19	-83.62	-28.32	-17.71	-16.33	-10.84	-29.30	-35.06	
Std.dev	14.31	28.84	1.17	6.35	3.58	22.73	26.92	35.53	20.46	24.62	33.03	
n	20	19	13	8	18	23	30	23	31	106	224	
t-stat	34.97 ***	-10.92	-0.35	0.17	-15.04	-0.60	-0.36	-22.04 **	-29.49 ***	-12.25	-15.88	
t+2												
Mean	1.33	0.00	13.35	-2.53	-16.70	0.00	-19.19	-17.57	-14.10	-0.83		
Std.dev	31.41	18.31	36.00	17.86	2.36	14.93	4.31	25.35	28.57	34.91		
n	20	19	13	8	18	23	30	23	31	106		
t-stat	0.19	0.20	13.36	-40.06 ***	-0.30	-0.27	-24.41 **	-33.23 ***	-27.48 ***	-24.35 **		

Table 23: Model 3: Excess returns: Governance. This table presents the results from Model 3. The tests are one sample t-tests. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%.

Building on these results and the recognition of the non-zero returns in isolation, a comparison between two groups and their relative performance will provide complementing insights. This will initially be approached by estimating whether there are differences in the average returns of companies engaged (including all engagements as one sample) and those firms not engaged in a particular period. As addressed earlier, the standard deviation within the sub-samples is also of interest and will be presented in this section.

All engagements applied on this model between the periods 2005-2016 generates a total of twelve independent t-tests. The number of observations in the test group ranges from a minimum of 51 in 2008 to the maximum of 529 in 2016. This ensures that the estimates will be reliable and accurate in line with the arguments of the central limit theorem (n>30). The number of observations can therefore be concluded as satisfying for testing hypothesis 17a and b for all periods. In seven of the twelve periods, there is a significant t-statistic result that is above the two-tailed critical level on 10% (approximately 1,65). This implies that there exists a significant difference in the returns between the test-group and the control-group in the periods 2007-2008, 2010-2011, 2013-2014 and 2016. The seven significant periods almost unanimously report a negative sign, except for 2007 which was positive. The periods that did not prove significantly different average returns between the two groups

were to a large extent negative estimates except for one year that was positive, giving a total of ten negative periods and two positive. These negative results conclude that engaged companies actually performed worse compared to those not engaged in terms of excess returns in that period, rejecting the null hypothesis of hypothesis 17a.

With that in mind, turning to the second area of investigating: the standard deviations. The periods where the test-group had a significantly different excess return proved lower standard deviations for the engaged companies compared to the non-engaged, except for the year of 2007 with an exceptionally high standard deviation of 126.3%. If broadening the scope by also looking at the results from the non-significant periods, the engaged companies had a lower standard deviation in all periods, except for in 2015 where it was slightly higher. To conclude the year-on-year results: The engaged companies tended on average have lower returns compared to the control-group but at the same time show signs of lower volatility, meaning they were less risky.

Pushing the previous test a bit further is to see if there exists lagged effects from *All* types of engagements. Since most of the engagements have been held privately and probably not been publicly available until next reporting period, it would be natural if the effects are lagged to some degree. To see if this holds true, the tests have been conducted in the following way: companies that have been engaged in a certain year, let's say in 2014, was compared to the non-engaged by looking at the excess returns in the subsequent periods, i.e 2015 and 2016. The number of observations in the test-groups ranges from a minimum of 51 (2008) to the maximum of 363 (2015), i.e the central limit theorem holds. The observed t-statistics can reject the null hypothesis of hypothesis 17b in seven of the eleven tests, just as many times as the previous test. There is a slight difference in the t-statistics compared to the previous test results. The significant results seem stronger, indicating that there is larger difference between the two groups. Just as the previous test, there are negative signs in six of the seven periods.

Again, it can be concluded that the results prove a difference between the two groups over the majority of periods tested and once again, the engaged companies seem to underperform. The standard deviations are also showing the same pattern as in the first test, showing a lower volatility among the engaged companies except for in 2015. Noticeable in 2015, the t-statistic is positive, showing signs of positive correlation between risk and return.

The last test on *All* engagements will be looking for any observable differences in excess returns two years after the engagements. This is of particular interest since one important aspect of being an active owner engaging in ESG-related issues is to enhancing returns and minimizing risk on a long-term perspective. Having the excess return lagged two years, there is no possibility to test the engagements in 2015 nor 2016. With that limitation, the last engagements are tested in 2014, with the excess returns controlled for in 2016. In total there are ten independent t-tests. Controlling for the number of observations in the test-groups, the lowest number of observations are 51 (2008) and a maximum of 158 (2014). Seven periods indicate a significant difference between the returns and all of them are negative, implying that the engaged companies had lower average returns. This again rejects the null hypothesis in hypothesis 17b as there is an evident difference between the test-group and control-group in the two years following an engagement. In terms of the standard deviations, all the periods show a lower volatility for the engaged companies, which is a positive sign when taking a risk-perspective. To conclude this last test on *All* engagements, experiencing effects two years following the engagements, the engaged companies are still performing worse but with lower risk in terms of volatility.

All Engagements

Engagement Year	**2005**		**2006**		**2007**		**2008**		**2009**		**2010**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	2.60	7.88	15.05	11.64	20.42	4.59	-5.29	1.68	14.83	30.62	-5.65	5.46
Std.dev	**21.4**	68.40	**27.2**	40.00	126.30	42.20	**16.10**	20.50	**68.50**	94.30	**22.50**	33.60
n	62	766	66	784	60	810	51	839	73	825	69	839
t-stat	-0.60		0.67		2.25 **		-2.37 **		-1.39		-2.68 ***	
t+1												
Mean	10.81	11.99	4.50	5.78	-4.43	1.70	14.88	30.21	-8.97	5.82	-2.55	-2.39
Std.dev	**22.6**	40.2	**32.5**	53.8	**14.6**	20.7	**46.4**	94.5	**30.5**	33.0	**22.0**	24.2
n	62	788	66	804	60	830	51	847	74	834	69	853
t-stat	-0.22		-0.18		-2.25 **		-1.14		-3.70 ***		-0.05	
t+2												
Mean	0.64	6.07	-3.64	1.68	4.56	31.11	-10.93	5.54	-9.70	-1.76	2.71	4.96
Std.dev	**32.7**	53.7	**15.2**	20.7	**48.5**	94.7	**20.4**	33.4	**22.4**	24.0	**23.9**	30.2
n	62	808	66	824	60	838	51	857	74	848	69	864
t-stat	-0.78		-2.04 **		-2.15 **		-3.47 ***		-2.73 ***		-0.60	

Engagement Year	2011		2012		2013		2014		2015		2016	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	-7.97	-1.82	3.86	4.88	-14.07	0.63	-11.54	-0.32	-3.73	-2.06	-3.51	0.77
Std.dev	**22.40**	24.10	**29.60**	29.80	**38.30**	43.50	**21.50**	31.20	30.10	26.90	**30.70**	42.90
n	87	835	77	856	80	861	158	797	362	610	529	448
t-stat	-2.27 **		-0.28		-2.91 ***		-4.31 ***		-0.89		-1.80 *	
t+1												
Mean	-4.59	5.76	-17.86	0.92	-11.79	-1.28	-7.49	-1.74	2.02	-3.65		
Std.dev	**25.9**	30.0	**38.7**	43.4	**24.3**	30.5	**26.9**	28.3	45.2	30.7		
n	87	846	77	864	81	874	159	813	363	614		
t-stat	-3.09 ***		-3.66 ***		-3.01 ***		-2.35 **		2.32 **			
t+2												
Mean	-15.31	0.88	-13.76	-1.16	-13.20	-1.72	-0.72	-1.71				
Std.dev	45.1	42.8	**24.4**	30.4	**26.9**	28.1	41.5	35.9				
n	87	854	77	878	81	891	159	818				
t-stat	-3.33 ***		-3.5 3 ***		-3.52 ***		0.30					

Table 24: Model 3: Excess returns Test vs Control – All engagements. This table presents the results from Model 3. The test group consists of all firms engaged during that given year. The control group contains all firms engaged at some point in the time interval 2005-2016 but not in the given year. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. Bolded Std.dev indicates that the standard deviation is lower for the test group compared to the control group in that year.

The next part of this model investigates the ESG sub-components; Environmental, Social and Governance. These tests will be conducted in the same manner as previous tests in this model. Being the pillars when categorizing the engagements in the data set, it is natural to isolate them and investigate if there are any differences between each component. The isolation will also provide better insights into their own specific effects and contribution to eventual differences in excess returns.

Environmental engagements are first out and by removing social, governance and any multi-engagements within same period reduces the dataset, leaving a minimum of observations on n=20 (2005) to the maximum of n=42 (2011) with most periods averaging around n=30. When the returns are not lagged, five periods prove to be significant at the 10% confidence level. The five periods are all showing a negative result except for in 2007. These results are following same patterns observed in the previous tests, even though these estimates are of slightly greater magnitude. Including the insignificant periods, most periods are of negative signs with few exceptions. This proves that when looking at the environmental sub-sample individually, the ability to reject the null hypothesis of hypothesis 17a holds for five of the twelve periods. The test-group only had higher standard deviations in three periods. One of the periods being 2007 where the standard deviation was remarkably high on

173%. A standard deviation of that magnitude was also observed in *All* engagements within the same year (2007), where it was 126.3%. Noticeable in 2007, with extremely high standard deviations, the average excess returns of the two groups are greatly positive, just like the result from *All* engagements. Taken together, it seems that the test-group is consistently proving to have lower standard deviation in the returns and that this holds for the environmental test-groups as well.

Estimating the effects of environmental engagements with one and two years lagged returns, the results are fairly similar to previous observations. The results from one year lagged effects also have five statistically significant results; 2007, 2009 and 2012-2014, albeit, this time all estimates are negative. In the non-significant periods, three estimates are positive with the rest being negative. Thus for hypothesis 17b there are some cases where the null hypothesis can be rejected but this is less than half of the instances so the tests fail to reject the null for the majority of the sample. The standard deviations from this test falls in line with the other results observed earlier. Here, the test-groups have lower standard deviations in the majority of the periods, with the only difference that the standard deviations for the test-group are rising above the control-group in the most recent periods, starting in the year of 2013. When lagging the effects on excess returns two years after the initial engagement, the majority of the tests proved a significant difference. Just like before, the estimates are negative with the exception of the last period, engagements being done in 2014 with excess returns tested in 2016. These tests further reject the null hypothesis in 17b that the lagged returns are not different between the two groups. The standard deviations are again favoring test-group for the majority of periods. Concluding the tests on how companies engaged on an environmental topic have performed in terms of excess returns in relation to non-engaged firms, the majority of periods show that they underperformed in relation to the non-engaged firms. This holds when the comparison is being made on a year-on-year basis and when looking at lagged effects from one to two years. Like the previous tests, the standard deviation is in favor of the engaged companies.

Environmental

Engagement Year	**2005**		**2006**		**2007**		**2008**		**2009**		**2010**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	4.16	7.57	20.56	11.57	35.09	4.60	-8.47	1.62	18.24	29.76	-7.71	4.95
Std.dev	**24.3**	66.8	**32.0**	39.4	172.9	41.9	**13.5**	20.5	**49.8**	93.8	**27.6**	33.1
n	20	808	32	818	31	839	30	860	33	865	24	884
t-stat	-0.22		1.27		3.18***		-2.67***		-0.70		-1.85*	
t+1												
Mean	12.87	11.88	13.12	5.40	-5.30	1.52	13.26	29.89	-9.72	5.17	-6.12	-2.30
Std.dev	**26.4**	39.4	**35.6**	53.0	**16.4**	20.5	**27.5**	93.9	33.0	32.9	**22.6**	24.0
n	20	830	32	838	31	859	30	868	34	874	24	898
t-stat	0.11		0.81		-1.83*		-0.96		-2.58***		-0.76	
t+2												
Mean	10.62	5.57	-6.96	1.59	9.20	30.06	-8.62	5.06	-11.06	-2.07	4.42	4.81
Std.dev	**37.9**	52.8	**14.8**	20.5	**56.7**	93.5	**19.5**	33.3	**22.5**	24.0	31.6	29.8
n	20	850	32	858	31	867	30	878	34	888	24	909
t-stat	0.42		-2.33**		-1.23		-2.23**		-2.14**		-0.06	

Engagement Year	**2011**		**2012**		**2013**		**2014**		**2015**		**2016**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	-4.64	-2.29	-1.28	5.00	-10.41	-0.32	-15.61	-1.86	-14.82	-2.20	2.06	-1.67
Std.dev	**22.3**	24.1	**27.4**	29.9	44.2	43.2	**18.9**	30.3	28.2	28.1	**35.7**	36.9
n	42	880	31	902	28	913	22	933	37	935	32	945
t-stat	-0.61		-1.15		-1.21		-2.11**		-2.67***		0.56	
t+1												
Mean	-2.07	5.12	-15.91	-0.10	-11.99	-1.86	-17.07	-2.34	4.53	-1.78		
Std.dev	**23.5**	30.0	**34.4**	43.5	30.8	30.1	33.1	28.0	68.9	35.1		
n	42	891	31	910	29	926	22	950	37	940		
t-stat	-1.52		-2.00**		-1.78*		-2.42**		1.02			
t+2												
Mean	-7.99	-0.27	-11.89	-1.85	-12.61	-2.37	27.04	-2.20				
Std.dev	**41.3**	43.4	**18.7**	30.4	29.1	28.1	61.5	35.9				
n	42	899	31	924	29	943	22	955				
t-stat	-1.12		-1.82*		-1.92*		3.69***					

Table 25: Model 3: Excess returns Test vs Control – Environmental. This table presents the results from Model 3. The test group consists of all firms engaged during that given year. The control group contains all firms engaged at some point in the time interval 2005-2016 but not in the given year. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. Bolded Std.dev indicates that the standard deviation is lower for the test group compared to the control group in that year.

Social engagements' effect on excess returns has been tested in the same way as environmental was above. The number of engagements in this category has a strong increase with time, especially in recent periods were the increase is rather drastic. For the year-on-year tests, the number of observations in the test-group ranges from the minimum of 13 (2008) to the maximum of 101 (2015). Testing a total of twelve periods, only one proves to be of statistical significance. This could stem from the fact that most years contained less than 30 observations leading to a lower accuracy of the model. Negative estimates proved to be the most common in the years where there was no significant difference between the two groups. These results are somewhat unique so far, since the previous tests have had several significant years when testing on twelve periods. This entails that for social engagements, the model fails to reject the null hypothesis of hypothesis 17a for practically all tests. When it comes to the standard deviations, the group of companies being engaged in the given time period had lower volatility for most of the periods. In 2009 the test-group experienced a standard deviation of 107% with the control group on a lower yet high figure of 92%.

When testing for lagged effects one year following the initial engagement, the results are quite different compared to year-on-year. From these eleven tests, a total of six periods prove to be of statistical significance. When companies have been socially engaged within the more recent years, starting from 2009 and onwards, their average returns proved worse than before, with the exception for 2010 and 2015 where it was positive. The social component thus proves to be better able to reject the null hypothesis posed in hypothesis 17b than what was the case for the environmental sub-sample, with a majority of the tests proving a significant difference. When the returns are lagged for one year, no extreme outliers appear in the standard deviations. Again, test-groups are less volatile in most periods, meaning they can be considered less risky to invest in from a shareholder's perspective.

Lagging the returns two years after the social engagement renders ten independent tests. These tests follow the results of one year lag rather closely. By eliminating the engagements in 2015, a large pool of observations in that test-group naturally disappears. The number of observations in the test-groups in the other periods are to be considered as low. With an average of only 20 observations per year, the accuracy of the tests is reduced. Yet, half of the tests prove statistically significant at the 10% level. All of these five estimates have a negative sign, meaning that two years after the social engagement took place, the test-group performed worse compared to the control-group. Most of the significant results are observed in the more recent periods. The test-groups are less risky speaking in terms of standard

deviation, following the same pattern as for the observations with one year lag. Here there are no extreme outliers and in some periods the test-groups are marginally riskier than the control-groups.

Taken together the results from this sub-sample aligns with previous tests in that the null hypothesis of 17b can be rejected for about half of the tests and a lower volatility tends to apply for a majority of years. Testing the isolated effects on the average excess returns made by social engagements year-on-year and with lagged years, the results prove fewer significant estimates for the year-on-year tests in comparison to previous samples. There seems to be a non-existent direct effect from social engagements, instead there appears to be significant difference looking at one and two years afterwards, mostly implying a lower return from the test-group than from the control-group.

Social

Engagement Year	**2005**		**2006**		**2007**		**2008**		**2009**		**2010**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	-3.30	7.78	8.64	11.97	-1.29	5.81	-4.35	1.37	10.24	29.82	-5.32	4.86
Std.dev	**21.3**	66.8	**17.5**	39.4	**30.6**	52.8	**19.5**	20.4	106.9	92.2	**20.6**	33.3
n	22	806	15	835	16	854	13	877	22	876	22	886
t-stat	-0.77		-0.32		-0.53		-1		-0.97		-1.42	
t+1												
Mean	8.60	12.00	0.98	5.77	-5.41	1.41	25.35	29.40	-8.32	4.93	1.64	-2.50
Std.dev	**25.7**	39.5	**25.6**	52.8	**13.6**	20.4	**67.6**	92.9	**32.8**	33.0	**20.9**	24.1
n	22	828	15	855	16	874	13	885	22	886	22	900
t-stat	-0.4		-0.34		-1.32		-0.15		-1.85*		0.79	
t+2												
Mean	-9.05	6.07	-2.23	1.34	-11.59	30.08	-7.43	4.79	-14.18	-2.11	4.55	4.80
Std.dev	**26.9**	52.9	**10.0**	20.5	**37.6**	93.1	**21.8**	33.2	**20.7**	24.0	**22.3**	30.0
n	22	848	15	875	16	882	13	895	22	900	22	911
t-stat	-1.33		-0.67		-1.78*		-1.32		-2.33**		-0.03	

Engagement Year	2011		2012		2013		2014		2015		2016	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	-6.40	-2.33	1.62	4.88	-14.85	-0.29	-1.58	-2.19	-8.44	-2.01	-5.36	-1.18
Std.dev	**22.3**	24.0	35.8	29.6	**28.8**	43.5	**18.7**	30.4	35.3	27.2	**28.5**	37.6
n	15	907	23	910	21	920	31	924	101	871	86	891
t-stat	-0.64		-0.51		-1.52		0.11		-2.17**		-1	
t+1												
Mean	-17.27	5.16	-22.02	-0.08	-12.90	-1.93	-16.29	-2.23	13.24	-3.27		
Std.dev	**27.9**	29.7	47.6	43.1	**20.0**	30.3	**26.8**	28.1	55.1	33.7		
n	15	918	23	918	21	934	31	941	102	875		
t-stat	-2.89**		-2.4**		-1.64*		-2.74***		4.31***			
t+2												
Mean	-28.02	-0.17	-12.49	-1.92	-12.69	-2.46	5.38	-1.77				
Std.dev	57.3	42.9	30.2	30.1	**22.1**	28.2	37.1	36.9				
n	15	926	23	932	21	951	31	946				
t-stat	-2.47**		-1.66*		-1.64*		1.06					

Table 26: Model 3: Excess returns Test vs Control –Social. This table presents the results from Model 3. The test group consists of all firms engaged during that given year. The control group contains all firms engaged at some point in the time interval 2005-2016 but not in the given year. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. Bolded Std.dev indicates that the standard deviation is lower for the test group compared to the control group in that year.

Governance is the last pillar of the ESG-components to be tested and its relation to the excess returns will be investigated in the same manner as with the environmental and social sub-samples. Starting with testing for direct effects from engagements made within the same period, the number of observations in the test-groups are fairly low for most of the periods, being below the desired number of n=30. Albeit, they do increase with time and just like the observations in social they escalate rapidly in the most recent periods, going from 31 in 2013 to 105 in 2014 and keeps on almost doubling itself from thereon to the maximum of 411 observed in 2016. The twelve tests generated four statistically significant periods with three being negative. Relating these results to hypothesis 17a, there are only a few results that can reject the null hypothesis and for the majority of the years the null cannot be rejected as there cannot be proved to be a significant difference in the returns between the two groups. Focusing on the tests that could not prove a significant difference at the 10% level, it should be noted that four periods have positive estimates which is more than what has been observed in earlier tests from this model. The test-groups have a lower standard deviation across all periods, meaning they bear less risk from an investor's perspective and there are no extreme observations, providing a uniquely stable return.

When estimating the lagged effects of the governance engagements one year afterwards, the number of significant observations falls to half, leaving only two periods with significant estimates and both of them are negative. Looking at the non-significant periods, they all reflect negative effects. This is definitely a noticeable change compared to the year-on-year test, where there were four positive and four negative estimates, even though the non-significance make the results less implicative. Once again and just like the year-on-year test, the standard deviation for the test-groups are continuously lower in all periods within this test, compared to the control-groups. With this, the null hypothesis that there is a difference between the two groups fails to be rejected for practically all tested years.

The last tests on governance engagements within this model, will be investigating if the engagements had any long-term effects on excess returns for the engaged firms. Testing ten independent samples, five of the estimates proved to be of statistical significance. The majority of the significant years are negative, meaning that governance engagements have led to lower returns on average for the firms over time. This means that the results are able to reject the null hypothesis of hypothesis 17b in half of the years where in the other half the null cannot be rejected. When testing the effects two years after an engagement, the year of 2015 has to be excluded since it is not possible today (read 2017), to validate the returns in a manner consistent with previous years. This leaves out a large bulk of observations, and leaves the rest of the periods with a lower number of observations. The standard deviations are once again in favor for the test-groups in almost all periods, except for the engagements that took place in 2013 and measured in 2015.

The results given from testing the effects on the average excess returns from firms engaged on governance topics differs somewhat in comparison with ones observed from testing the same model on environmental and social engagements. In governance there are to some extent more positive results, even though the majority of them is not of statistical significance at the 10% level. The second difference stems from the standard deviations as they are almost exclusively in favor for the test-group across the different periods, with just one exception.

Governance

Engagement Year	**2005**		**2006**		**2007**		**2008**		**2009**		**2010**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	7.53	7.48	10.84	11.93	12.16	5.59	5.11	1.25	14.17	29.65	-3.82	4.83
Std.dev	**17.7**	66.8	**23.5**	39.4	**35.3**	52.7	**16.2**	20.4	**32.0**	93.3	**18.7**	33.3
n	20	808	19	831	13	857	8	882	18	880	23	885
t-stat	0		-0.11		0.44		0.53		-0.7		-1.23	
t+1												
Mean	11.19	11.92	-7.23	5.97	-1.14	1.32	3.92	29.57	-8.36	4.87	-2.83	-2.39
Std.dev	**14.3**	39.6	**28.8**	52.9	**11.6**	20.5	**63.5**	92.8	**23.5**	33.2	**22.7**	24.0
n	20	830	19	851	13	877	8	890	18	890	23	899
t-stat	-0.08		-1.08		-0.43		-0.77		-1.68*		-0.08	
t+2												
Mean	1.33	5.79	0.84	1.29	13.35	29.57	-25.29	4.88	-1.67	-2.41	-0.84	4.94
Std.dev	**31.4**	52.9	**18.3**	20.4	**36.0**	93.1	**17.8**	33.0	**23.6**	24.0	**14.9**	30.1
n	20	850	19	871	13	885	8	900	18	904	23	910
t-stat	-0.37		-0.09		-0.62		-2.57***		0.12		-0.91	

Engagement Year	**2011**		**2012**		**2013**		**2014**		**2015**		**2016**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	-13.41	-2.03	13.03	4.59	-16.85	-0.07	-13.62	-0.76	0.23	-3.55	-3.55	-0.09
Std.dev	**22.3**	24.0	**24.5**	29.9	**39.1**	43.3	**22.0**	30.7	**27.2**	28.4	**30.8**	40.7
n	30	892	23	910	31	910	105	850	224	748	411	566
t-stat	-2.55**		1.34		-2.12**		-4.15***		1.76*		-1.44	
t+1												
Mean	-1.77	5.01	-16.33	-0.22	-10.84	-1.88	-2.93	-2.65	-3.51	-0.96		
Std.dev	**26.9**	29.9	**35.5**	43.4	**20.4**	30.4	**24.6**	28.6	**33.0**	38.0		
n	30	903	23	918	31	924	106	866	224	753		
t-stat	-1.22		-1.76*		-1.62		-0.09		-0.9			
t+2												
Mean	-19.19	-0.01	-17.57	-1.79	-14.10	-2.30	-8.26	-0.73				
Std.dev	**43.0**	43.2	**25.3**	30.1	28.5	28.1	**34.9**	37.0				
n	30	911	23	932	31	941	106	871				
t-stat	-2.39**		-2.48**		-2.29**		-1.98**					

Table 27: Model 3: Excess returns Test vs Control –Governance. This table presents the results from Model 3. The test group consists of all firms engaged during that given year. The control group contains all firms engaged at some point in the time interval 2005-2016 but not in the given year. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. Bolded Std.dev indicates that the standard deviation is lower for the test group compared to the control group in that year.

Achieved engagements is last sub-sample tested in this model. Having tested the effects of *All* engagements, with results to the most part suggesting that average returns for the test-groups are significantly lower than the average returns for the non-engaged control-group, further investigating this relationship is of great interest. This is approached by investigating the effects on excess returns from those engagements being classified as *Achieved*. In this definition of the model, the control group is thus constructed by both engaged and non-engaged firms, but with the important distinction that engaged firms in the control group have not achieved an objective in the given year. As described earlier in the text, worth mentioning again, engagements classified as achieved and almost achieved might be easy to define in theory but harder to categorize in practice as the level of accomplishment can be very similar in some cases and rather different in other. This non-linearity in the definition of the variables might affect the outcome of the tests but attempts to reduce such influences are likely harming the reliability of the study more than it benefits the results. Thus, with the definitions intact, isolating the achieved engagements will bring further clarity and understanding of the potential effects successful engagements can have on a company's performance.

The tests follow the same approach as previously, i.e. the first test is estimating the difference in average excess returns the same year as the initial engagement for companies with an *Achieved* objective compared to a control group. This will be followed by testing for any lagged effects conducted as previously. Having narrowed the dataset, the sample size in the test-groups are smaller compared to the tests on *All* engagements. The sample size range between n=11 observations (2008 & 2013) to a maximum of n=75 (2015), where most periods have observations around n=23. Since most of these number are <30, the implications of the central limit theorem weaken, also weakening the assumption of normally distributed data and thus creating wider confidence intervals in the estimations.

The results show coherence with the previous test but prove a weaker statistical strength, in line with the expectations. When the returns are not lagged, three out of twelve observation years prove statistically significant, the years of 2013, 2014 and 2016 and they are all negative. The reduced number of significant estimates could either be related to inefficiencies in the estimations in lines with what has been discussed above. Alternatively, the fact that the engagements have in fact been achieved indicates that the negative results from the full sample model are to a large extent due to engagements that has not been achieved. Taken together, these tests affirm the negative relation between the engaged firms and excess returns, however testing the achieved engagements signals a smaller negative effect than when including also non-achieved engagements. This entails that the null hypothesis of

hypothesis 17a can only be rejected for a few years and that the tests to a large extent fails to reject the null hypothesis. The insight that the test-group has a lower standard deviation holds for this sub-sample as well. This is impressive seeing as the test-group theoretically is much less diversified than the control group due to the low number of observations and thus a higher standard deviation from the test-group could have been expected.

When estimating the average excess return of an *Achieved* engagement one and two years after the initial engagement date, the results are similar to the case with no lags used. The results suggest that a firm who was engaged in 2012 with an achieved objective did not experience a significantly negative direct impact on its excess returns in 2012, but instead in 2013. Firms engaged in 2013 did experience heavy struggles according to these tests. It shows that these firms, although having achieved their engagement objectives, on average experienced lower excess returns in year 2013, 2014 and 2015 thus indicating a strong inability to turn ESG initiatives into value enhancing activities in the eyes of the shareholders. This effect was similar when looking at *All* engagements[12] in the test-group as well. For the firms engaged in 2014 the effect is not as strong, yet the engaged firms prove to have a significantly lower excess return compared to the control group both in the year of the engagement but also one year after. Engagements done during 2008 also seemed to have more long term effects as they experienced a significantly negative effect in 2010. This provides further motivation to reject the null hypothesis of hypothesis 17b as the differences across the two groups hold also for periods t+1 and t+2. The standard deviation remains lower for the test group in most of the years and through all three model specifications and serves as one favorable feature of the engagements.

[12] Please refer to appendix 9.4 for a graphical presentation on the comparison between All and Achieved engagements excess returns and standard deviations

Achieved												
Engagement Year	**2005**		**2006**		**2007**		**2008**		**2009**		**2010**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	1.50	7.66	11.09	11.93	3.10	5.73	2.52	1.27	15.03	29.66	0.44	4.70
Std.dev	**22.8**	66.9	**24.6**	39.5	**27.0**	52.9	**18.0**	20.4	**53.8**	93.2	**24.3**	33.2
n	23	805	27	823	17	853	11	879	20	878	18	890
t-stat	-0.44		-0.10		-0.20		0.20		-0.69		-0.53	
t+1												
Mean	10.16	11.96	4.55	5.72	-2.65	1.36	12.97	29.54	0.12	4.72	-5.88	-2.33
Std.dev	**22.8**	39.5	**33.3**	53.0	**9.5**	20.5	**37.9**	93.0	**30.8**	33.1	24.2	24.0
n	23	827	27	843	17	873	11	887	21	887	18	904
t-stat	-0.21		-0.11		-0.80		-0.58		-0.62		-0.62	
t+2												
Mean	3.74	5.74	-4.88	1.48	7.57	29.76	-12.64	4.82	-10.64	-2.21	2.00	4.85
Std.dev	**36.1**	52.9	**12.4**	20.5	**61.6**	93.0	**28.2**	33.1	25.4	23.9	**21.1**	29.9
n	23	847	27	863	17	881	11	897	21	901	18	915
t-stat	-0.17		-1.59		-0.97		-1.74*		-1.59		-0.40	

Engagement Year	**2011**		**2012**		**2013**		**2014**		**2015**		**2016**	
Group	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control	Test	Control
Year-on-Year												
Mean	-2.95	-2.39	-2.32	4.97	-30.31	-0.27	-12.47	-1.91	-5.37	-2.45	-21.11	-1.05
Std.dev	**18.3**	24.1	34.6	29.7	**37.3**	43.2	**23.7**	30.2	29.3	28.1	**21.2**	37.1
n	20	902	22	911	11	930	24	931	75	897	24	953
t-stat	-0.10		-1.13		-2.29**		-1.69*		-0.85		-2.63***	
t+1												
Mean	0.56	4.89	-22.54	-0.09	-21.52	-1.95	-12.74	-2.42	-6.88	-1.09		
Std.dev	**19.5**	30.0	**31.8**	43.4	**19.5**	30.2	**25.7**	28.2	38.5	36.7		
n	20	913	22	919	11	944	24	948	76	901		
t-stat	-0.64		-2.40**		-2.14**		-1.77*		-1.31			
t+2												
Mean	-20.74	-0.18	-19.91	-1.75	-30.31	-2.36	6.37	-1.74				
Std.dev	**24.4**	43.5	**26.5**	30.1	**23.6**	28.1	58.3	36.2				
n	20	921	22	933	11	961	24	953				
t-stat	-2.10**		-2.80***		-3.28***		1.06					

Table 28: Model 3: Excess returns Test vs Control –Achieved. This table presents the results from Model 3. The test group consists of all firms engaged during that given year. The control group contains all firms engaged at some point in the time interval 2005-2016 but not in the given year. The asterix notation indicates the level of significance where; * = 10%, ** = 5% and *** = 1%. Bolded Std.dev indicates that the standard deviation is lower for the test group compared to the control group in that year

6. Analysis, Discussion and Implications for Schroders

In this section the results from chapter five will be analyzed and related to the insights presented from previous research and theoretical frameworks introduced in chapter two. The purpose of this chapter is to bring the results to a more general level and discuss the implications this has for Schroders, other investors and the topic of ESG-investment. The analysis will be conducted following the structure of the models as above but taken together and analyzed as a whole in the later part of the chapter. As mentioned in chapter three, the direction of causality is central in interpreting the results. This is not explicitly addressed in the methodology of the paper and due to the extent of such explorations, separate tests have not been conducted to establish the direction of specific relations. Instead this will be addressed continuously in the analysis below, allowing the reader to assess what seems most plausible but also encouraging further research to solidify the direction of the causality.

6.1 Objective result

The results from model 1 enables a discussion around how successful Schroders have been in achieving the engagements initiated and what types of firms might be the easiest or hardest to change. As identified by Deutsche Bank (2015), it is probable that there are varying effects depending on whether the firm engages in ESG as a combined topic and whether the components E, S and G are handled individually. This is immediately recognized with this data set as well, as all three sub-components prove different in their relation to the success of the engagement. Before further emphasizing the specific results, it should be recognized that achieving an engagement depends on several factors and it has not been possible to address all quantitatively in the model. One important aspect is how easily achieved the objective is. It is likely the case that some suggested objectives are more easily implemented and achieved both between the individual ESG-components but perhaps also within each one. Applying this perspective to the results, one would argue that environmental engagements seem to be the "easiest" ones to achieve and perhaps these types of objectives require a consistently lower amount of change, effort, time, investment, commitment or similar. This would also indicate that governance engagements are the hardest to achieve and thus they would be the most complex ones. Whether this is true or not is ultimately defined by Schroders as they have the level of insight needed to deem whether one ESG-component generally has lower requirements for the engagement objectives, however it seems unlikely.

This opens up for a second perspective to interpret the results from. If the complexity is assumed the same across the ESG-components, the willingness of firms to implement changes on certain ESG-

areas could differ. This would suggest that firms are more eager to implement environmental changes, compared to the other two components, and that there would be one or several traits of this that firms value highly. In accordance with Endrikat (2016), environmental activities of firms can have both positive and negative effects on the stock price, insinuating the idea that it is important for managers to adhere to the market standards to initially avoid any negative effects and ideally also reap some benefits. Such arguments support the notion that managers would be very willing to initiate environmental activities proposed by their investors as they can be a source of improved stock market performance. Such attractive features are arguably also seen in social activities. Being a rather well-explored area with the highest capital allocation amongst the three components with money manager funds (Deutsche Bank, 2015) and with a long history of research coverage (Margolis et al., 2007), the area of CSR surely contains attractive features for managers if interacted with appropriately. Nonetheless, such a high level of standardization could make it less attractive for managers to fully engage with CSR objectives as it could arguably be seen more as mandatory than value adding and unique, ultimately neutralizing the urge for such initiatives. Furthermore, the relation between CSR and CFP is versatile and thus financial managers might have issues isolating the benefits and have experienced average results previously, making them less eager to drive social initiatives more than "necessary". This idea finds support in the findings from Margolis et al. (2007) as they identified an overall effect that was positive yet small and that more than half of the research done fails to find significant relations between the two. Adding to this, social activities might have the ability to preserve CFP rather than to generate it (Godfrey et al., 2009). Together these aspects might be something that managers consider and subsequently categorize CSR-activities more precautionary than value adding which in turn might dampen their eager and motivation to fulfill social engagements, yet keep them to a moderate extent.

Noting that governance engagements are negatively related to the probability of achieving an engagement seems both evident yet contradictory. In adopting changes to the governance system, the managers are more often than not affecting their own influence, role and importance within the firm and in the relation with investors and other stakeholders, something that usually is not seen as favorable. There are also arguments suggesting that investors perceive changes or adjustments to the governance system as a negative signal and that it communicates negative information, making it unattractive for managers to engage with such issues as it might cause short term damage. A contradiction arises within the fact that large bodies of research have documented a positive relation between CFP and governance which entails that such activities can be value adding and thus should

be attractive to managers. Similar to the area of CSR, governance structures have been widely investigated and especially from a stakeholder perspective but likewise from the shareholder perspective of the firm; a well-constructed governance system can be a powerful tool in aligning interests across the organization and its stakeholders. Arguing that managers would be inclined to achieve such favorable results and thus should be willing to commit fully to changes in this area could be supported by these insights. The negative relation is opposing this and signals that one or two other effects are present. Either governance initiatives are harder to change as they affect the organizational structure more incrementally and thus only fractions or parts of them are achieved, or managers stand more unwilling to implement such changes as they might not improve their position within the organization. With two strongly opposing effects, the fact that governance initiatives have proven a high volatility in its relation to CFP as suggested by previous research, this report affirms that there are some unattractive features within governance engagements. (Deutsche Bank, 2015).

These observations prove to hold also when investigating differences in engagement topics. One of three topics related to environmental topics is positively related to the probability of achieving an engagement whereas the social topics do not prove to have an effect different from zero. Interestingly, there are some governance topics that have a significantly positive influence on achieving the objective while one is identified as negative. The fact that some governance topics are positive, with the estimate of the whole sample of governance engagements being negative, clearly proves that some objectives are more easily achieved than others within governance. The Business Integrity and Corporate Strategy topics, that to some extent both relate to the internal characteristics of the firm and how the firm chooses to operate, are part of the responsibilities of the manager(s) in the firm and thus aligns well (and does not threaten) with their position. On the other hand, Shareholder rights issues are arguably more separated from the everyday operations of the firm and can be seen more of a separated issue. If looking at it in the crassest way, increased shareholder rights means less power for the managers. Naturally, it would take a lot more effort and likely higher levels of support to implement changes related to such a topic compared to the other two. In highlighting this, the most important insight is that although governance as a single sample was estimated to have a significantly negative effect on the probability of achieving an engagement objective, this effect is driven by specific items and that there are topics and characteristics within governance that are more easily implemented and thus attractive also in the eyes of the managers. This finding, with support from for example Gillian & Starks (2000), indicates that finding engagement topics that aligns with the interest of managers can be detrimental in how successful the engagement will be.

These patterns remain, and are further strengthened, for different sectors within the ESG-components as well. It turns out that some sectors have a greater or lower ability to achieve their engagements. The Telecom Services and Utilities show a higher probability of success in their engagements than any other sector where the IT sector proves to have the most negative effect on succeeding with an objective. Investigating structural differences between the sectors might provide further insight into what types of firms are more inclined to fulfill engagement objectives. From the descriptive statistics provided above there are some differences worth highlighting between the most positive sectors and the most negative one. Beginning with ROA the Telecom and Utilities sector have lower average return on assets than the sample as a whole whereas the IT sector proves to have a ROA a bit higher than the sample average. Combining this with the fact that the Telecom and Utilities sectors also have more assets than average, and the IT sector lower assets than average, there seems to be a clear difference between the sectors. The more capital intensive industries with lower returns on their assets have to a larger extent achieved their engagement objectives. There are several reasons as to why this might be the case. One could be that they wish to increase their returns given the same assets (what firm would not want that) and that they find ESG-activities as potent ways of achieving this. It could be through less waste, more resourceful machinery, higher customer loyalty or better aligned managerial incentives that does not necessarily alter the assets of the firm but could increase the return. Alternatively, it could be that firms with higher ROA already have reached a more satisfying level of internal efficiency and consequently are not experiencing similar pressure to improve their performance, making them less eager to initiate ESG-activities. Another interesting insight stemming from the descriptive statistics is that IT firms on average have a higher Tobin's Q and M/B-ratio compared to both the sample average and the average for the two most positively affected firms. Together those figures open up for the interpretation that a higher market valuation might add pressure to the managers of the firm. This possibly makes the managers more inclined to have a shareholder perspective of the firm rather than incorporating other stakeholders, ultimately leading to them not committing fully to ESG-initiatives. These implications are teasing, as they clearly identify certain characteristics that have been proven significant in achieving engagement objectives. However, they should be considered with caution. Looking at the other sectors that proved significantly different from zero but with a lower magnitude, the relations mentioned above becomes less consistent. Whether ROA or M/B-ratio are the correct measures in determining the likelihood of any given firm achieving an objective can be questioned. Nonetheless, it should stand clear that there are differences between sectors in how well the managers can incorporate ESG-initiatives and that the discussion above potentially identifies some of these.

It is not only the engagement topic and the sector of the firm that proves to be important but also how the firms are engaged. Collaborative engagements are proven to be less efficient in accomplishing the given objective. Traits of collaborative engagements that could potentially cause this can be multiple and perhaps not well aligned agendas of the different actors, greater inefficiency in the processes as more people involved does not necessarily mean better and faster results, greater complexity of the objectives agreed upon and greater levels bureaucracy. Not only does the likelihood of succeeding with an objective increase when Schroders is the sole engager, there is also a reward for repeatedly engaging with the same firm. Having a larger number of engagements, it is likely that the two parties are more familiarized with one another and are able to cooperate more efficiently and mutually beneficially which is reflected by the positive effect this has on the success of engagements. Notable is that both effects are solely driven by governance engagements and the effect is not present for the other ESG-components.

As noticed in the previous paragraph, some of the firm characteristics that are primarily to be considered as control variables in the models have significant relations to the probability of an engagement being achieved. These will not be analyzed to a great extent yet there are two variables that are deemed interesting to bring up for further discussion. Contrasting what was mentioned above, the size of a company has a negative impact on the probability of achieving an engagement. As argued in chapter four, the relation between firm size and performance variables such as ROA and Tobin's Q is well documented but the relation to a firm's willingness to achieve objectives has much less documentation. The model used is able to conclude that a positive relation between firm size and the probability of achieving an objective does not exist as there is a negative effect present for three of the four samples tested. In isolation this might not seem surprising, a larger firm might have more rigorous structures, a more complex organization with long term contracts and capital investments already in place. A smaller firm is normally seen as more agile, faster to adapt and more dynamic, making it a better target and more likely to successfully implement changes to the operations. On the other hand, as argued above, with greater size comes greater resources and leeway to perform ESG-initiatives which also could be backed by greater publicity. Nonetheless, the former effect was proven to be the case for this data set. Interpreting this becomes more complex when also realizing that the effect of ROA is significantly positive for the entire sample. There is a disconnection between the three variables. Firm size seems to have a positive impact on ROA as proved by earlier studies, a higher ROA increases the probability of successfully implementing an ESG-objective as proved in

this model but firm size in itself is negatively related to the probability of success in an engagement. This implies that there is something inherent in the return on assets, not attributable to the magnitude of assets that is driving the positive influence on achieving engagements. This immediately turns the focus to the return part of the ROA variable which is ultimately determined by the sales. The sales growth proves to be significantly positive in this model which implies that it is not necessarily the size of the firm that is helpful, but that if targeting firms that are growing in terms of revenue, one increases the chances of seeing the objective achieved. Intuitively, such a connection seems plausible as a growing firm is constantly adding to their revenues and thus growing organically. In turn, this often means changes and additions to the organization of some sort, aspects that could make it easier and more attractive for managers to add ESG-initiatives.

Having shown sensitivity to various sorts of engagement topics, sectors and firm characteristics, the probability of success also varies greatly across regions. As clear as the results are, it is equally difficult to identify the root cause for the different effects of various geographical regions. The EMEA region is the only region that positively influences the probability of success where the remaining regions except for South America are significantly negative. To dwell on macroeconomic factors that might cause these differences is outside the scope of this paper. One fact that could contribute to these results is that seen to the number of engagements, UK is the most common country and Europe in general is quite heavily represented. With Schroders being a UK-based firm, their local knowledge and expertise might be better translated in their engagement efforts in this market where they could potentially hold greater experience and expertise and thus be more credible as compared to in other markets. The contribution of this geographical segment of the test is best seen when juxtaposing the EMEA region with a significantly positive estimate to the North America region with a significantly negative estimate. Two arguably similar markets with opposing results. To understand why takes more detailed knowledge about Schroders operations in the two regions. To prove a difference between them and to highlight that the outcome of engagements is to a non-negligible level dependent on the geographic setting is nonetheless an important contribution.

More of illustrative purpose than of great implication for Schroders, composing the type of engagement that is most likely to succeed is an enticing exercise. Based on the results from this model, the optimal engagement objective (counted as the sum of probabilities of being achieved) would be on the governance topic of corporate strategy, have had engagements done previously with the same firm, targeting a firm within the EMEA region that is operating within the utilities sector and has a

high ROA, sales growth and high R&D spending in relation to their assets. Surely, there are plenty of firms matching this description and although aggressively targeting those firms is not suggested the conclusion from this section that all these factors do play an important role in the likelihood to achieve an objective is applicable to the future operations of Schroders.

6.2 Corporate Financial Performance

In model 1 ROA, Tobin's Q and Sales per employee were estimated to check for relations to the probability of achieving an engagement. Model two adds to the understanding of these variables by testing whether these effects go the other way as well. For example, ROA was determined to have a positive relation to the probability of success however when seeing if engagements have a similar influence on ROA, the significance vanishes. With result estimates varying around zero, there seems to be little to no evidence that an engagement leads to improved return on assets either in the short or the medium term. What this indicates is that there is some aspect of ROA that is helpful when wanting to follow through on engagement objectives, yet the objectives themselves cannot be said to improve a firm's return on assets.

As mentioned in chapter four, ROA is a good measure for comparison within the same sector or region but might not prove as useful when looking across a broader, in this case global, sample. It could be that the cross-regional and cross-sector aspects of ROA are averaging each other's effect towards zero in the full sample and that the vast number of variables potentially affecting ROA makes the model incapable of capturing them specifically. The panel data model was chosen as it efficiently controls for omitted variables like this and there might be other reasons as to why ROA does not prove to be affected by engagements. One could be that Schroders are investing in medium to large firms both in terms of revenue and assets meaning that for a variable to have a significant effect on ROA it requires improvements of great magnitude. Because of this the model also investigates whether there are any improvements, defining any improvement in ROA regardless of magnitude as the same. The relationship becomes somewhat more distinct but remains inconclusive. What can be deducted from the results on ROA is that the effects are strongest and the most stable one year after the engagement, indicating that the effects are not immediate but that it could potentially be able to realize gains (or losses) in a short period of time. An important aspect in interpreting a ratio like ROA is that it can remain constant although the variables are changing. It would be an interesting exercise to investigate how the revenues are changing with regards to ESG-activities and to assess whether there is a significant relation between the two as that would help understanding if the engagements

are able to generate additional revenues for the firms. Likewise, it could be that an eventual increase in revenue has necessitated an increase in assets as well, holding ROA constant. By taking this model one step further, future studies could generate deeper insights to this area. If finding that the revenues does actually increase, assuming everything else constant, Schroders and their target firms could more specifically address how to achieve the same results without also increasing the asset base. If the returns would prove unchanged or negative, it would instead be important for Schroders to find ways to create a connection between the ESG-activities and revenues rather than focus on how to efficiently deploy the assets.

The fact that the models are unable to prove a difference in the effect between the achieved engagement objectives and all engagements is significant in its insignificance. That there is no additional reward for actually implementing/changing/introducing ESG-initiatives compared to having a discussion about the topic or the target firm not making any efforts to meet the objective, is surprising. It could be expected that the estimated coefficients would be more positive and also of greater magnitude for the achieved engagements reflecting the additional efforts in those cases. However, no such effect is visible. The immediate interpretation is that the engagements are not value-enhancing. Such a conclusion is however extremely premature and it can be out ruled by the fact that in themselves, each engagement is in fact a positive and value enhancing thing. In recognizing this, it is more important to assess whether the efforts are actually measurable as that is the purpose of the model. The fact that engagements differ in their influence on ROA, makes it plausible that there is a discrepancy between engagements in their inherent nature that averages out the positive influence of some with a non-quantifiable effect of others.

The neutral effect that engagements have had on ROA applies to Tobin's Q as well and a discrepancy between the engagement and dependent variable can be identified in much the same way as above. In measuring the value of "off the book items", it could have been the case that engagements fail to produce the real increases in returns but that they create other favorable traits with the engaged firms that are valued by investors. However, such an effect is not documented either and although some of the results are stronger than others, the overall effect seems to be insignificant. There are two ways of approaching such a result. One is to look at what the engagements change and argue whether or not that could be considered valuable for investors. The other way is to look at what the investors base their valuations from and how that relates to the ESG-activities undertaken. The former approach has been covered by various theoretical and empirical works and although this has to no extent been

covered explicitly or in full here, there is a general consensus that there are activities and ways in which addressing ESG-issues increases the value of the firm. The second approach relates largely to the discussion of pricing externalities introduced in chapter two and briefly brought up in the discussion of ROA above. Even though it is a bold assumption and one that might not hold in all the situations, assuming that if done appropriately, an ESG-engagement is value enhancing for the firm, then these results communicate a disconnect between the given activity and the valuation of it which. If keeping this assumption for the purpose of reasoning, the inability for engagements to be reflected in Tobin's Q is motivated by the inability to measure their contribution which is in turn dependent on common tools and practices of valuing off balance sheet activities and ultimately their contributions to the bottom line. In line with this, one can note that the test estimates in the environmental sample proved to be lower and less positive compared to the social and governance samples. The paper by Mattison et al. (2011) highlights that the great discrepancies that exists within the pricing of environmental externalities created by a firm, and thus also rewards for good actions, is a great market failure and it affirms the weak relation that exists between environmental engagements and market valuations. Further backing this argument is the fact that there are both positively and negatively significant variables relating to Tobin's Q so the market does react to the engagements, but not to the desirable extent. Relaxing the assumption that all ESG-engagements are inherently value adding and instead questioning whether the engagements actually are value adding, it is encouraged to take these results one level deeper and to use more specific variables in the model. It could be measuring the waste reductions done through environmental activities, or by finding a tool to value customer loyalty increases in relation to social initiatives or to quantify how much shareholders are willing to pay for greater shareholder rights. What this model can contribute with in understanding this is the fact that the achievement of the objectives did not alter the effects. In a sense, that indicates that the engagements and their effects are not reaching the broad public. As neither a negative or positive relation is detected, it could be that the private trait of the engagements is hindering the information to reach the market. In a scenario when that is the case, there will not be a difference between an achieved engagement objective and an engagement discussion as neither of them will be communicated to the investors in a way that enables them to adjust their valuations of the firm accordingly.

Not finding positive effects from initiating ESG-activities in the variable Sales per employee, the connection between corporate responsibility and employee productivity is not prevalent in this data set. For ESG-activities to have the desired effects for the employees, there are some aspects that are

of central importance. Firstly, the initiative must be aligned with the values of the organization. This is more easily said than done as it not only requires the ESG-project to be somewhat tailored to the organization, but also that the firm is well aware of their internal values and that these are consistent throughout the organization. If there is a misalignment between these, it is plausible to argue that the employees will not experience motivational increases from the project and thereby not exercising greater productivity as hoped. A second aspect is how well the ESG-activity is considered to be an integral part of the organization's core activities rather than a separate project. If there are only a few employees that are working with and in relation to the ESG-initiative and the influences it might have on the firm, then only a few employees will change their behavior. For an engagement to have an effect throughout the organization, and for it to prove a significant improvement in the employee productivity, the actions cannot be separated from the majority of the organization. Thirdly, if the ESG-engagement is not handled appropriately, say ended prematurely or redirected to a less coherent area, it could spur the opposite reaction and rather create dissatisfaction amongst the employees. Taken together these arguments point to the notion that in order to increase the Sales per employee, an engagement needs to be tailored to the individual organization and that the insignificant results prove a lack of customization from Schroders and their engagement partner. Having no additional effect from actually achieving the engagement objective, there is a notion that trying is equally valuable as succeeding with ESG-initiatives which could indicate that the more engagements and ESG-activities a firm initiate (but not necessarily completes), the greater influence this has on the employees. Such a conclusion is not possible to draw based on this paper, however the results here provide insights that could support it.

In general, the results are to a large extent insignificant, not proving a relation between ESG-engagements and the CFP variables chosen. This affirms an important realization that has been suggested by some previous research as well and it provides an indication of where the ESG-topic as a concept is at. In the theoretical review in chapter two numerous studies that measured highly specific parts of the different E, S & G categories were presented, many of which have been summarized by the Deutsche Bank (2015) report. Such specific studies seem to, more often than not, find significant relations to various CFP-measures but reports estimating the relation from an ESG-perspective has a lower ability to identify these relations. Expressed differently, specific and narrow efforts within individual parts of the ESG-umbrella seem to be easier to measure, value and thus evaluate compared to broad, all-encompassing efforts. This is motivated by the great difference in effects across engagement definitions. Such a realization is also intuitively compelling as the more different effects

brought into an analysis, the more normalized the effect becomes and thus on average the inclusion of more engagement definitions removes the overall effect. Although statistically this concept is basic, the implications for the purpose of this thesis are more complex. Suggesting that there is no significant influence of ESG-engagements on CFP variables, the results of this model are arguably confirming this line of reasoning.

This turns the discussion into why there are differences between certain types and characteristics of the engagements and why the effect of these is possible to capture in some cases and in some cases not. Generally, such analysis is hard given the foundation of this paper but there are some traits that can be highlighted as plausible determinants. The most obvious reason could be that the ability to value and the level of standardization greatly varies not only within E, S or G but also within geographic regions, engagement topics, sectors etc. Having inconsistencies in this area disconnects the engagement value chain and harms the ability to conduct tests and analysis on several areas simultaneously. Some areas, for example an engagement suggesting a way to reduce waste in a production process, are easily observable, measured, quantified and integrated with existing tools and reporting structures (i.e. is easy to include in a balance-sheet). Contrastingly, for example a social engagement suggesting the implementation of targets and commitments to improved employee health & safety, some engagements are drastically harder (if at all possible) to observe which in turn makes them close to impossible to incorporate in company reports in a consistent and standardized way. These differences might well be the reason for the insignificant results experienced within this model. A second trait that has been addressed throughout the paper is the level of achievement and as discussed above, the disconnect between effort and reward that seems to be present in this data sample is reducing the incentives for firms to actually go through with an engagement objective. Either because the benefits are realized in an early stage or because the benefits are not graspable.

The discussion above has initiated some reflection on whether there simply is no relation or if there are some disconnections between the activities and the performance measures used. This discussion is interesting and it offers some reassuring interpretations; The ESG-engagements done by Schroders are not dispersing resources, managerial focus or valuable assets and there through reducing the performance of firms. Taken together the results are disproving any arguments saying that engagements are conflicting with the core business of the firm, distracting managers, allocating resources inefficiently or any similar unwanted scenario. In this light, the firms can rejoice as performing good and doing good does not conflict with one another. This is important in aligning the

firms and the managers as they are engaged in situations where they might be in doubt or skeptical about the process. Contrasting this effect slightly is the fact that the models could not prove an increased reward or benefit from increased efforts. This makes it harder, but more important, to construct the engagements in a manner that applies to each individual organization and that aligns well with their internal development so that the achievement of an objective is seen as natural and integral and not as unrelated and a burden. Nonetheless, the engagements cannot be proven to add value to the organization either. For Schroders such a gap in the relationship can be dubious as it disconnects their efforts from an eventual reward which is troublesome. What can be said on that note, based on the discussion provided here, is that one could analyze the process from the end to the start, rather than from beginning to finish. Thus one could investigate the driving factors behind improvements in, say, ROA and what activities have positively influenced this. Taking these activities and attempting to apply ESG-initiatives to those can prove to be an efficient way of leveraging the already successful activities with further improvements and value adding components.

6.3 Excess return

The results from this model rather unambiguously communicates that the firms engaged in a given year produce negative returns both in terms of their expected return and in absolute values. This contradicts earlier research where the relation to CFP (broadly defined, CFP also captures the stock returns) has been proven positive, yet the interpretations of the result differ slightly. Most studies to date are estimating the relation based on public data, meaning that their studies are relating observations at a later stage in the engagement value chain to CFP. Contrastingly, this study in company with a few more is based on private information and consequently investigates if this relation holds also from earlier steps. Dimson et al. (2015) conducts similar tests but with a more sophisticated model that is testing the ESG-returns in relation to a group of comparable firms, not engaged or chosen by any ESG-criterion. Their ability to prove positive CARs indicates that there are valuable aspects in ESG-engagements that the model applied here is unable to capture. Investigating what this could be is of equal interest as the conclusion itself. What is striking with the results from the first tests in this model is that there are mostly negative results but that they do not seem to change over time. In the year of the engagement, a negative average return is experienced in many of the measured years yet this effect seems to last over the two subsequent periods as well. It becomes even more interesting when realizing that the effect is not the same in the three periods but it is changing somewhat in the years with significance but with similar interpretations. It seems like the market has a rather long memory when it comes to ESG-initiatives. This is the case for all three ESG-components

and it brings up the idea that an engagement that leads to an observable action (from the market) is priced not only instantly but also at least for two years.

Although this effect does not follow any clear patterns, the existence of one indicates that the market is observing the initiatives and also incorporates it in their valuations. What could drive this effect is that the engagements done by Schroders do not contain value adding components and thus they are seen to destroy or neutrally affect shareholder value. It could also be the other way around that Schroders are, intentionally or not, engaging poorly performing firms with the notion that adding the ESG-component to their agenda will add value. The negative to neutral market reaction could in this case be an indication that the ESG-initiatives have not been able to improve the value of the firm's activities but that it instead remains on a path of poor to negative performance. The method of this paper differentiates itself from the one of Dimson et al. (2015) for example as it includes tests that are comparing engaged firms' excess returns in a given year to other engaged firms instead of a set of industry peers or similar. This allows for interpretation around the relation just described. By seeing how a group of companies perform in relation to a group of companies that have been targeted using the same criteria, the difference that the engagement creates is somewhat isolated.

Investigating the tables for the excess return models, it is clear that the underperformance of engaged firms holds in more settings. Separating the firms into two groups and controlling for differences in average returns across them, an important finding, similar to one of Dimson et al. (2015), is that achieving engagements is important for the stock price and the value it can create to investors. The average excess return of the test-group is notably improved when the engagement has been classified as achieved and the effect holds through all three periods tested. For the sake of clarity, the returns of the most recent years are still to the most part negative and the improvement is seen in relation to the tests on all engagements where the test-groups significantly underperformed the control-group. This is not the case when Schroders, in cooperation with the portfolio company, manages to implement the assigned objective. With this finding, the model underscores the importance of achieving objectives but more implicitly, this also signals the importance of setting objectives rather than having discussions/meetings or similar conversational engagements that do not incorporate any subsequent actions. Such a conclusion might seem trivial, yet it also suggests that having defined objectives that are related to ESG-activities might be more efficient and value-creating than having management realize the potential value and trying to incorporate it more discreetly into its operations.

The observant reader might realize that it is probably not the definition "achieved" that is improving the average return for the engaged companies and that setting engagement objectives at a basic and low level in order for the objective to be classified as successful is likely not the way forward. Instead, relating back to model 1, realizing that certain firms are better approached in certain ways and the idea that having objectives that are mutually appealing both for Schroders and for the organization is not only theoretically sound but now also factually so. The benefits of achieving an engagement could be defined as containing two separate effects; one communication effect and one action effect. The communication effect relates to what information is shared with the public and the value of this. An example of a negative communication effect could be that Schroders representatives meets with managers at the portfolio company and for some reason this becomes publicly known. Such information could create worries amongst investors that something is not right and the stock price could be struck with greater skepticism. A positive communication effect could come from the firm being able to introduce new initiatives, projects or activities and by sharing this, the firm creates good news for investors, affecting the stock price positively. Distinguishing the communication effect and the action effect is that the communication effect does not by definition require an action. An action effect however, can only be realized through an activity. More specifically it entails activities that in any way (measureable or theoretical) can affect the firm and its operations. This effect can also be positive or negative. A negative action effect would be engaging in a project targeting ESG-activities that are not expected to be valuable or that are irrelevant. It could also come from un-achieved activities. Not being able to finish a project or successfully undertake new tasks is not seen as favorable and failing to implement, or abandoning, an objective could be very harmful for the firm, both in terms of stock market reactions but also reputation wise. A positive action effect is simply realized when the firm undertakes actions that are relevant, value-adding and successful. With these two effects, a discussion engagement can only generate a positive or negative communication effect and the lack of tangible action removes any effects from such. It is apparent that a firm and the engagement-activist want to achieve both a communication effect and an action effect and to the best of their ability tailor the engagement so that both these turn out positive. That way the firm is successfully improving its operations and/or performance and at the same time managing to communicate this efficiently, in the end being more likely to see favorable stock price movements.

Although there are clear incentives to conduct objective engagements and have them achieved, it is still of relevance to assess why other engagements generally receive negative reactions from the market. One aspect that is introduced by previous research and made additionally clear from this paper

is the fact that looking at ESG as a concept rather than analyzing the components individually harms the results and the ability to draw insights from the research. In the results from this model it is evident that the effects seen for the whole sample is driven by one or two of the ESG-components but all three rarely have comparable effects. By looking at the effects of engagements jointly is thus removing the individual effects and noting that the areas include a great variety of topics that are all very different from each other. The year of 2014 is illustrating this effect as the joint sample proves a significantly worse performance for the test-group but where the social sub-sample independently was estimated to outperform its control-group. Comparably, in 2015 the difference between the average returns of the two groups is insignificant when the components are tested jointly. Individually, there are significant differences between the test-group and the control-group for all three components but where the environment and social sub-samples were underperforming and the governance test-group over performed in relation to their control-groups.

When the excess returns of the engaged companies are tested individually in accordance to their ESG-component, the negative returns remain. If applying the results from Barnett and Salomon (2006), it implies that it is not necessarily the engagements themselves that are not valuable to investors but rather the construction of such a portfolio that is not favorable. They reason that portfolios with low screens are able to perform well as they can enjoy the full benefits of solid diversification and are not exposed to any idiosyncratic risk. Similarly, portfolios that are constructed based on high levels of screening are able to perform above average as the screens are efficiently identifying firms performing better than the average firm and so the portfolio is constructed on a smaller universe of firms but a universe with higher quality. The portfolios created in this model, the test-group and the control-group, only have two screens; The first one is applied by Schroders and is determined by the fact that the firm has been engaged on an ESG-topic in the time span of 2005-2016. The level of this screen is unobservable given the data and information provided, creating aspects of uncertainty. The second screen is whether or not the firm actually was engaged in the given year or not. Introduced in chapter three, there are some levels of screening done from Schroders before an engagement is made, making the first scenario with no or low levels of screening not likely present here. Whether the screening process from Schroders is sufficient to capture the benefits of high levels of screens and that this can offset the loss from imperfect diversification in the portfolio, remains unsaid. Nonetheless, the presence of negative to neutral average returns for the test-group could be related to this and is subsequently indicating that either applying more thorough screening or removing existing ones could improve the stock market performance of engaged companies.

If further investigating reasons why the engaged companies are unable to generate positive returns in a consistent manner, it could be interesting to evaluate the materiality of the engagements. The point here relates to the findings of Koutsantonis et al. (2016) and the important distinction between material and immaterial ESG-projects. As an investor, companies with numerous positive NPV-projects in their pipeline are attractive and appropriately rewarded as these projects are seen to improve the value of the firm. In much the same way, as an ESG-activist you would value firms that have multiple topics that could be addressed and improved so that there are many value-enhancing projects rather than just a few or none. The point being that by engaging or investing in firms with "good" ESG-standards, the number of value adding and material projects will be lower and instead the engagements might address immaterial aspects within ESG. Immaterial engagements are arguably more socially popular, but are also more likely to result in unfavorable stock market reactions as they tend to lack a financial impact. Relating back to the discussion about the tradeoffs between a high or low screening process, reasoning that having low screens can be seen as additionally attractive. Partly because it makes it possible to enjoy the full benefits of diversification as stated above but also because the ability to initiate more material and value adding ESG-projects seems favorable.

If accepting lower, or even negative returns, one would expect to be compensated in other areas. Seeing that the standard deviation of the test-groups is rather consistently lower than that of the control-groups, the risk-reward relation seems to hold well within the sample. This risk-reduction component inherent in the engaged firms offers benefits to both theorists and practitioners. It strengthens the notion that responsible activities have insurance-like abilities and subsequently that they might be more efficient in preserving market value than creating it as proposed by Godfrey et al. (2009). It also suggests that ESG-activities can protect, or even immunize, companies to large market reactions related to certain events by reducing the risk of negative events internally but also proactively managing relations so that harmful market events can be more easily handled. As implied in these results, investors put a more long term credibility to the firm with the belief that ESG-initiatives signals sound business decision making and consequently also a higher belief in the future performance of the firm.

Setting the scene for studies to build on these results, the expected returns applied in the calculation of the excess returns relies on the capability of CAPM. The skeptical reader will note that the ability to prove excess returns significantly different from zero stems from poorly estimated expected returns

and claim the CAPM inappropriate to use in such settings. This analysis is not going into detail within this area but it encourages more research to conduct similar tests with different methods of calculating expected returns to complement this paper and enrich the topic further.

6.4 General insights and Implications for Schroders

By having engaged companies on a great number of different topics, in a global setting with no observed preferences for different sectors, firm traits or similar, Schroders have approached the ESG-topic with an all-capturing approach. Naturally this leads to a great variety of problems, issues, question but also rewards, benefits and insights. This has been proved in this paper and the tests conducted have been able to identify differences in how engagement traits are accepted within the target firm's organization. The implications of this goes beyond the fact that achieving set out objectives is an accomplishment in itself.

Aligned with the findings of Dimson et al. (2015), the stock market reaction is more favorable for achieved engagements than when looking at all engagements together. When testing the full sample (i.e. including all engagements) this paper has made the case that engagements are reflected negatively in stock returns more often than the opposite but where about half the results are indistinguishable from zero. One important contribution and similarity to comparable papers is the improved stock performance that is identified when studying achieved engagements. In affirming and strengthening this suggested relation, achieving the engagement is argued to be of great importance for Schroders as their performance ultimately is related to the stock price. A short-sighted and narrow suggestion would be to make the objectives basic and easily achieved so that this effect is enjoyed. Such naïve approach is likely insufficient and instead there are some other attributes that achieving an engagement brings to the organization and the value creation. The achieved engagements are likely those that align well with the firm's core operations and thus is easy yet powerful in the sense that the objective resonates with values at the firm yet has a significant impact on the operations.

In joining these insights, it entails that investors do recognize some value creation from a firm achieving an objective. It could be that achieved objectives are communicated differently, that they are constructed as to be done exhaustively and not in bits and pieces for them to work or that they are defined in a way that accentuates synergies and efficiency gains when the objective is completed. Generally, the idea that it is easier to do something if you have succeeded at something similar applies well. The relevance of this increases when considering the test results from the CFP models as they

could not prove improvements when the engagements are achieved over when all engagements are tested. Thus, investors value something with achieved engagement objectives that is not visible within the performance measures chosen.

From an optimistic point of view where the objectives have an effect on the underlying variables investigated, the objective must have increased or decreased both variables in similar magnitude within each of the ratios included in CFP. The objective could have increased the return of a given firm but it might also have increased the assets needed to generate such a return, leaving the ratio unchanged. This is an inherent issue when analyzing ratios, identifying in which part of the ratio the change is taking place, and it is an area where further research should be done in order to determine the relationship so that any subsequent actions could be done with more detailed knowledge. Taking a more skeptical standpoint, the lack of significance and improvement in CFP from achieving an objective can be motivated by the fact that there is no improvement at all. It has been argued previously but highlighted once more here, such a skeptical standpoint is not taken in this paper as it does not resonate with the purpose of it nor with previous research.

Taken together, the engagement value chain that has been identified by this paper can be seen as a process starting with the engagement and ending with the stock price reaction, drawn out as in figure nine. It has been duly noted that when estimating the effect from all engagements, there is no significant effect in CFP and in general a negative to zero change in the stock price. However, when trimming the data to the achieved engagements only, the stock market reaction turns to a neutral reaction and, here comes the point, it is not through an increase in the CFP measures. Seeing an improvement in CFP would motivate a more favorable stock market reaction but as this relation has not been proven there are other, unidentified benefits from the achieved objectives that the investors are valuing.

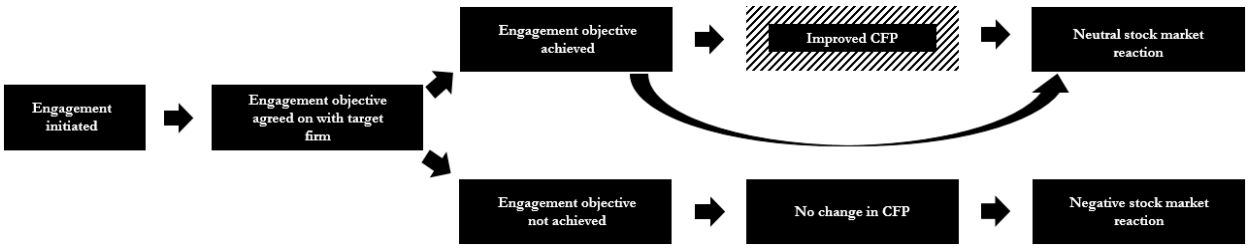


Figure 9: The ESG-engagement value chain

The most appealing motivation for this is that the objectives are generating benefits that are not reflected in the CFP measures. It could be that they apply to other metrics, are only visible on a longer time frame, or have benefits that are not measureable but easily communicated that catches the likings (or rather neutralizes) investors. Nonetheless, it seems plausible that if the engagements would also activate an improved CFP, it would be more accessible to the investors and they would be able to value it with less caution and uncertainty, something that is likely to also improve their financial prospects of the activities. To leverage this opportunity, Schroders needs to identify what effects the ESG-objectives have on the underlying variables of CFP (defining CFP more broadly than has been done in this paper) and focus more on *incorporating* the ESG-initiatives to existing resources and assets rather than *adding* to them. This suggestion is based on the idea that being able to relate ESG-activities to well-known and frequently used financial metrics will reduce the distance to investors to value the activities and the benefits stemming from such and in effect this completes the value chain.

One trait that investors are concerned about and value, arguably to a comparable extent as the returns, is the risk in a given investment or portfolio. When estimating the excess returns of the engaged firms the risk-reward relation holds rather well and the generally lower returns of the test-group is complemented with consistently lower standard deviations. Considering this from a portfolio perspective, this adds an attractive feature and it indicates that the firms in the test-group are less exposed to market movements as their returns move within a smaller interval. Whether this is because of the engagements or because Schroders are engaging firms that has low volatility remains somewhat undetermined however the fact that a group of engaged firms produce less risky returns is a highly appealing finding. Where the firms that had achieved certain objectives experienced improved market reactions compared to the engagements that had "only" been engaged, the standard deviation does not prove a similar pattern. The standard deviation for the tests-groups with the criteria of having had an achieved objective is somewhat closer to the standard deviation of the control group, yet the test-group presents a trend of lower standard deviations. Although it is to some extent opposing the importance of achieving the engagement objectives, this serves as an attractive trait. It signals that the risk reduction component that comes with the ESG-initiatives is inherent in the earlier stages of the initiatives but does not seem to be dependent on the success of them to the same extent as seen with the returns. Firms seem to be able to reduce the volatility of their stock price by showing interest in and to some extent initiating activities within ESG.

The somewhat unique perspective of this paper, where having the viewpoint of an institutional investor, makes these result highly insightful and rewarding. It complements much of the previous studies done and can provide Schroders with further implications. As risk is a largely standardized and well-used measure, it is another useful aspect to underline when evaluating the engagements. Being unable to prove higher performance in terms of stock returns, the fact that the model is able to distinguish the test-group from the control-group in another favorable way is positive. This implies that Schroders in their efforts could emphasize this characteristic and prove that ESG-engagements has a risk management component that if composed correctly can be of great value in a portfolio. Through this, not only can they immunize companies from large crisis like the one Volkswagen experienced recently but they are also able to include ESG-engagements as a detrimental aspect in their portfolios as it could ultimately improve their risk-adjusted returns.

Throughout the models, the ESG-components have been broken down individually and when possible, this has been related to results from the full sample. An apparent trend is that the relations presented from the full sample is often driven by one of the ESG-components. What is observed is that the effects differ between the components and that when grouping efforts together the effects disappears. This alleviates two separate but central implications. Firstly, the differences between ESG-components and other defining traits of engagements also means differences in how to observe, measure and value such activities. This is noted when looking at the difference between engagement discussions and engagement objectives. The objectives are observable as they incorporate a counter reaction from the firm, they are measurable as they can be quantified as done in this paper and they can be valued as various methods can assess their contribution to firm performance aspects. Discussions on the other hand are strictly informative and lacks these important features. When looking at the aggregated effects of these two, it seems likely that the effects to some extent takes out each other. Further, similar discrepancies exist between engagement topics as noted above and the inability to relate certain engagements to a performance metric is argued to harm the ability to prove the benefits from them. This implies that Schroders might be able to realize greater efficiency and effort-reward relation if the engagements are focused more to have objectives that are readily measurable and relatable to internal performance metrics so that their contribution becomes evident, accessible and easily communicated. This relates to the discussion held above where the materiality of the engagements is levitated as an important aspect in order to establish a clear connection to CFP. By identifying companies that have several material positive ESG-NPV projects available,

engagements could more closely affect CFP as they more intrinsically improve operational aspects of the firm. If also done in a more specific manner, isolating areas that are known to have a higher likelihood of success and that are easily quantified, improvements in CFP would become more observable and readily available for investors, ultimately with the aspiration of improving the share price. Thus, finding these firms and defining the specific engagement definitions that contain these favorable traits is proved to be central. This paper has targeted the second part of this relation and has affirmed the ability to find characteristics and aspects of engagements that are more important than others.

A second important implication is that the results of this report, in alignment with previous research, concludes that grouping the components together rather than keeping them separated can have harmful effects on the results. The motivation for including engagements and objectives on several ESG-components simultaneously is intuitively clear, the more you do the better. However, this is not proven to be the case. An illustrative example can be taken from table 27 where the results from the excess returns comparisons are presented. In the year 2015, the full sample test-group experienced significantly lower returns than the control-group and when the sample is broken into each individual ESG-component the environmental and social test-groups are also underperforming their respective control-groups but where the governance test-group is significantly outperforming its control-group. Taking a more individual and specific approach is further motivated by the notion that more focused efforts are more efficient. This said, it is not suggested that one of the three areas is given more relevance but rather that when engagements are made, there are potential benefits from conducting, measuring and evaluating engagements on a disaggregated level. Included in this is a requirement that Schroders, in coordination with the target firm, identifies suitable objectives and to some extent prioritize these in accordance with the organization's strategic position, goals and structures. Having a narrower scope on the objectives and by initially prioritizing certain types of engagement traits, Schroders could focus their efforts on engagements where the value chain above is intact and if also keeping distinctions clear between various engagements this value creation will enjoy both positive communication and action effects. The idea being that to form greater expertise in the areas where the market is developed enough to maintain the value chain so that when ESG as a concept matures and the tools and practices for such valuation improves, Schroders will be in a leading position to expand and address a wider scope of engagements. This way, Schroders aligns more with what is considered beneficial today yet puts themselves in a position so that when the pressure from investors and other stakeholders to more explicitly value ESG-activities, they can expand their expertise.

7. Conclusion and focus of future research

This study has been performed with the purpose to investigate a private data set provided by Schroders and to explore if their engagements with portfolio companies on ESG-related topics have been successful. As asset managers, the ultimate definition of success is the return on investment. The approach taken in this paper has been to not explicitly focus on the returns related to the engagements but to take a more holistic perspective in an attempt to explore areas that are affected by such activities. In effect, this complements existing research well and in the theoretical review in chapter two, the diverse and complex nature of responsible investments has been covered. The general consensus from this section is that, as of today (May, 2017), there is no clear consensus whether responsible investments, actions and initiatives are value-enhancing for the individual firms or for the investors. Many can prove a positive relation between the two, some prove a negative and some point to a neutral relation and so it is of interest to investigate one a more thorough level with a somewhat different approach. This is allowed for thanks to the nature of the data set and the paper has been able to study the process that takes place before it becomes publically available and measurable. This unique opportunity has also strengthened the motivation to have an exploratory approach in order to identify any relations between the ESG-engagements and firm performance, providing practical insights and additional perspectives when attempting to find causal relations.

The first effort to investigate the engagements was to assess what determines whether an engagement is successful or not. With the notion that succeeding with the engagement is the goal, isolating traits that have improved the probability to succeed with an engagement is important. This is studied using an ordered probit model where the results show some insightful aspects. There seem to be great differences between ESG-components, sectors, topics and regions that greatly influence the likelihood of a firm completing the suggested change. It highlights the fact that not all engagement definitions are seen as equally attractive and that, the preferences of the managers at the portfolio companies are extremely important to align with. As they will implement the suggested initiative, the importance to make the engagement mutually beneficial and that this is well communicated is proved by the results from this model. It is encouraging to further investigate what factors are increasing the willingness to undertake ESG-initiatives within a firm. One such area that is briefly mentioned in this study, are different firm specific characteristics. It is noted that sales growth (in terms of revenues) are positively affecting the probability to succeed with an engagement. To investigate more thoroughly what firm

characteristics are beneficial in this area is a focus of future studies that could continue to make engagements more successful.

The second model, and the second area of focus, was to estimate the relation to CFP. This study has defined CFP with the variables of ROA, Tobin's Q and Sales per employee as they offer great insights individually but also as they complement each other contributes jointly. To test this in a consistent and thorough manner, the data was fitted into a panel data model that allows for both time and firm fixed effects, minimizing the omitted variable bias and improving the estimates. Generally, no significant relation to the performance measures chosen can be proved, either in the year of the engagement or two years after. This turns into a discussion around whether the activities related to the engagements are observable, measurable and valuable and related to aspects brought up in chapter two, there seems to be a great discrepancy in the pricing and valuation of ESG-activities. The implications for Schroders are that in order to transfer the value from the activities to the balance sheets of the companies and ultimately to the stock market, this model and the accompanying discussion suggests that more focused efforts should be employed in areas where there are standardized measures that can capture the added value. By starting with a more narrow focus, the market demand and increased interest in responsible activities will drive the topic forward, expanding the number of ESG-topics that can be included as they become possible to measure and quantify. An obvious focus for future research is thus how to price and value ESG-activities within companies and how the benefits added to the company can be better reflected into the stock market valuations. Being a very broad topic, it captures aspects such as what investors prioritize and look at in their valuations, what is contributing to improved performance and why some activities might generate more value than others.

The last model in this paper focuses on this, the relation to the stock market, as it measures the ability to generate excess returns that the engaged companies have proven. The tests prove negative to neutral returns almost exclusively and further strengthens the implications from model two, that there is a disconnect between the inherent value of ESG-initiatives and the value they reflect in the market. The model is simple in its design, using t-tests to assess performance over the years to see whether engaged firms in a given year have performed better or worse than those firms not engaged in that year. One specification of the model shows that having achieved engagements are improving the returns and puts them around averages of zero. This finding is important as it underscores the importance of fulfilling the engagement objective, and a discussion around the materiality of the engagements is

central to such implications. Taken together, this model concretizes the findings from the previous models as it shows the importance of achieving the given engagement objective yet indicates that the efforts and initiatives are not positively rewarded. Such a realization is important from the perspective of Schroders as it signals an inability to create tangible value from the engagements. This paper indicates that as is, the active ownership from Schroders is to a certain extent inefficient and that in order to continue these activities in a way that is favorable and attractive to their investors, there needs to be further efforts in defining and specifically addressing areas where the engagement value chain is intact and the efforts can be rewarded appropriately. The result here contrasts some existing findings and it proves the diverse nature of studies like this one where different methods and approaches continue to yield different results. However, for future research and for Schroders internally to focus more on this relation it is highly important to approach the topic specifically and first establish consensus within for example the specific ESG-components and their relation to performance measures and then build the smaller pieces together instead of approaching the topic as one.

8. Bibliography

Accenture. (2016). *Insight: UN Global Compact CEO Study.* Retrieved from https://www.accenture.com/us-en/insight-un-global-compact-ceo-study

Agresti, A. (2002). *Categorical Data Analysis* (2 ed.). John Wiley & Sons.

Ashbaugh-Skaife, H., Collins, D. W., & Lafond, R. (2004). *Corporate Governance and the Cost of Equity Capital.* Retrieved from https://ssrn.com/abstract=639681

Ashbaugh-Skaife, H., Collins, D. W., & Lafond, R. (2006). The effects of corporate governance on firms' credit ratings. *Journal of Accounting and Economics, 42*, 203-243. doi:10.1016/j.jacceco.2006.02.003

Barnett, M. L., & Salomon, R. M. (2006, November). Beyond Dichotomy: The Curvilinear Relationship between Social Responsibility and Financial Performance. *Strategic Management Journal, 27*, pp. 1101-1122. Retrieved from : http://www.jstor.org/stable/20142399

Barth, M. E., Konchitchki, Y., & Landsman, W. R. (2013). Cost of capital and earnings transparency. *Journal of Accounting and Economics, 55*, 206-224. doi:10.1016/j.jacceco.2013.01.004

Bebchuk, L. A., Brav, A., & Jiang, W. (2015, June). The Long-Term Effects of Hedge Fund Activism. *NBER working paper series no*. doi:10.3386/w21227

Bebchuk, L., Cohen, A., & Ferrell, A. (2009, February). What Matters in Corporate Governance? *The Review of Financial Studies, 22*, 783-827. doi:10.1093/rfs/hhn099

Becht, M., Franks, J., Mayer, C., & Rossi, S. (2010). Returns to Shareholder Activism: Evidence from a Clinical Study of the Hermes UK Focus Fund. *The Review of Financial Studies*, 3093-3129. doi:10.1093/rfs/hhn054.ra

Bharadwaj, A. S., Bharadwaj, S. G., & Konsynski, B. R. (1999). Information Technology Effects on Firm Performance as Measured by Tobin's q. *Management Science, 45*, 1008-1024. Retrieved from http://dx.doi.org/10.1287/mnsc.45.7.1008

Bhojraj, S., & Sengupta, P. (2003). Effect of Corporate Governance on Bond Ratings and Yields: The Role of Institutional Investors and Outside Directors. *The Journal of Business, 76*, 455-475. doi:10.1086/344114

Bickel, P., Friedrich, R., Link, H., Stewart, L., & Nash, C. (2006). Introducing Environmental Externalities into Transport Pricing: Measurement and Implications. *Transport Reviews, 26*, pp. 389-415. doi:10.1080/01441640600602039

Black, B. S. (1998). Shareholder Activism and Corporate Governance in the United States. *The New Palgrave Dictionary of Economics and the Law, 3*, 459-465. Retrieved from https://ssrn.com/abstract=45100

Brammer, S., & Millington, A. (2008, December). Does It Pay to Be Different? An Analysis of the Relationship between Corporate Social and Financial Performance. *Strategic Management Journal, 29*, 1325-1343. doi:10.1002/smj.714

Bryman, A., & Bell, E. (2011). *Business research methods* (3 ed.). Oxford: Oxford University Press.

California Public Employees Retirement System. (2016, February). *CalPERS Investment Fund Values*. Retrieved from Investments: https://www.calpers.ca.gov/page/investments/asset-classes/asset-allocation-performance/investment-fund-values

Carleton, W. T., Nelson, J. M., & Weisbach, M. S. (1998, August). The Influence of Institutions on Corporate Governance through Private Negotiations: Evidence from TIAA-CREF. *Journal of Finance, 53*, 1335-1362. doi:10.1111/0022-1082.00055

Cox, T. (1993). *Cultural Diversity in Organizations: Theory, Research, and Practice*. Berrett-Koehler.

Cremers, K. J., & Nair, V. B. (2005, December). Governance Mechanisms and Equity Prices. *Journal of Finance, 60*, 2859-2894. doi:10.1111/j.1540-6261.2005.00819.x

Daykin, A. R., & Moffatt, P. G. (2002). Analyzing Ordered Responses: A Review of the Ordered Probit Model. *Understanding Statistics, 1*(3), 157-166. doi:10.1207/S15328031US0103_02

Del Guercio, D., & Tran, H. (2012). Institutional Investor Activism. In J. R. H. Kent Baker (Ed.), *Socially Responsible Finance and Investing: Financial Institutions, Corporations, Investors and Activists* (pp. 359-380). John Wiley & Sons. doi: 10.1002/9781118524015.ch19

Deutsche Bank Climate Change Advisors. (2012, June). *Sustainable Investing: Establishing Long-term Value and Performance.* Retrieved from https://www.db.com/cr/en/docs/Sustainable_Investing_2012.pdf

Deutsche Bank: Asset & Wealth Management. (December 2015). *ESG & Corporate Financial Performance: Mapping the global landscape.* Retrieved from https://institutional.deutscheam.com/content/_media/K15090_Academic_Insights_UK_EMEA_RZ_Online_151201_Final_(2).pdf

Dhaliwal, D., Li, O. Z., Tsang, A., & Yang, Y. G. (2014). Corporate social responsibility disclosure and the cost of equity capital: The roles of stakeholder orientation and financial transparency. *Journal of Accounting and Public Policy, 33*, 328-355. doi:10.1016/j.jaccpubpol.2014.04.006

Dimson, E., Karakas, O., & Li, X. (2015). Active Ownership. *The Review of Financial Studies, 28*, pp. 3225-3268. doi:10.1093/rfs/hhv044

Ding, H., Zhao, Q., An, Z., Xu, J., & Qian, L. (2015). Pricing strategy of environmental sustainable supply chain with internalizing externalities. *Int. J. Production Economics, 170*, pp. 563-575.

Edmans, A., Fang, V. W., & Zur, E. (2013). The Effect of Liquidity on Governance. *The Review of Financial Studies, 26*, 1443-1482. doi:10.1093/rfs/hht012

Endrikat, J. (2016). Market Reactions to Corporate Environmental Performance Related Events: A Meta-analytic Consolidation of the Empirical Evidence. *Journal of Business Ethics, 138*, pp. 535-548. doi:10.1007/s10551-015-2598-0

European Sustainable Investment Forum. (2016). *European SRI Study.* Retrieved from https://www.eurosif.org/sri-study-2016/

Flammer, C. (2015). Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach. *Management Science, 61*, 2549-2568. Retrieved from http://dx.doi.org/10.1287/mnsc.2014.2038

Freeman, E. R. (2010). *Strategic Management: A Stakeholder Approach* (25th anniversary edition ed.). Cambridge: Cambridge University Press.

French, K. R. (2017). *Current Research Returns*. Retrieved March 2017, from Kenneth R French: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Friedman, M. (2002). In *Capitalism and Freedom: Fortieth Anniversary Edition* (p. 133). Chicago: Univesity of Chicago Press.

Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance*, 2646-2654. doi:10.1016/j.jbankfin.2008.06.002

Gill, J., & Johnson, P. (2010). Theory and management control. In *Research Methods for Managers, 4th edition* (pp. 38-68). SAGE.

Gillan, S. L., & Starks, L. T. (2000). Corporate governance proposals and shareholder activism: the role of institutional investors. *Journal of Financial Economics, 57*, 275-305. doi:10.1016/S0304-405X(00)00058-1

Godfrey, P. C., Merrill, C. B., & Hansen, J. M. (2009, April). The relationship between corporate social responsibility and shareholder value: an empirical test of the risk management hypothesis. *Strategic Management Journal, 30*, 425-445. doi:10.1002/smj.750

Grewal, J., Serafeim, G., & Yoon, A. S. (2016, July). Shareholder Activism on Sustainability Issues. *HBS Working Paper Number: 17-003*. doi:http://dx.doi.org/10.2139/ssrn.2805512

Heal, G. (2005). Corporate Social Responsibility: An Economic and Financial Framework. *The International Association for the Study of Insurance Economics, 30*, pp. 387-409. doi:10.1057/palgrave.gpp.2510037

Investment & Pensions Europe. (2016). *World's largest Asset Managers 2016*. Retrieved March 2017, from https://hub.ipe.com/top-400/total-global-aum-table-2016/10007066.article

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics 3*, 305-360.

Jo, H., & Harjoto, M. (2011). Corporate Governance and Firm Value: The Impact of Corporate Social Responsibility. *Journal of Business Ethics, 103*(3), 351-383. doi:10.1007/s10551-011-0869-y

Jo, H., Kim, H., & Park, K. (2015). Corporate Environmental Responsibility and Firm Performance in the Financial Services Sector. *Journal of Business Ethics, 131*(2), 257-284. doi:10.1007/s10551-014-2276-7

Klassen, R. D., & Mclaughlin, C. P. (1996, August). The Impact of Environmental Management on Firm Performance. *Management Science, 42*, pp. 1199-1214.

Kotsantonis, S., Pinney, C., & Serafeim, G. (2016). ESG Integration in Investment Management: Myths and Realities. *Journal of Applied Corporate Finance, 28*(2), pp. 10-17.

Larcker, D. F., Richardson, S. A., & Tuna, İ. (2007, July). Corporate Governance, Accounting Outcomes and Organizational Performance. *The Accounting Review, 82*, 963-1008. Retrieved from http://www.jstor.org/stable/30243484

Margolis, J., Elfenbein, H., & Walsh, J. (2007). *Does it Pay to be Good? A Meta-analysis and Redirection of Research on the Relationship between Corporate Social and Financial Performance.* Working paper Harvard University.

Mattison, R., Trevitt, M., & van Ast, L. (2011). *Universal Ownership: Why environmental externalities matter to institutional investors.* Retrieved from http://www.unepfi.org/fileadmin/documents/universal_ownership_full.pdf

Mcwilliams, A., & Siegel, D. (2000, May). Corporate Social Responsibility and Financial Performance: Correlation or Misspecification? *Strategic Management Journal, 21*, 603-609.

Morgan Stanley Capital International. (n.d.). *ESG Integration*. Retrieved from https://www.msci.com/esg-integration

Pasta, D. J. (2009). Learning When to Be Discrete: Continuous vs. Categorical Predictors. *SAS Global Forum, 248*. Retrieved from http://support.sas.com/resources/papers/proceedings09/248-2009.pdf

Pfarrer, M. D. (2010). What is the purpose of the firm? : Shareholder and stakeholder theories. *Good business: exercising effective and ethical leadership - New York*, pp. 86-93.

Pratt, S. P., & Grabowski, R. J. (2014). *Cost of Capital: Applications and Examples* (5 ed.). Dawsonera: John Wiley & Sons Ltd.

Robson, C. (2013). *Real world research: A resource for users of social research methods in applied settings* (Vol. 3). Hoboken, N.J. ; Wiley.

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (Vol. 7). Harlow, England: Pearson Education Limited.

Schauten, M. B., & van Dijk, D. (2011). Corporate Governance and the Cost of Debt of Large European Firms. *ERIMREPORT SERIES RESEARCH IN MANAGEMENT*, 46. Retrieved from http://hdl.handle.net/1765/19679

Siew, R. Y., Balatbat, M. C., & Carmichael, D. G. (2016, April). The impact of ESG disclosures and institutional ownership on market information asymmetry. *Asia-Pacific Journal of Accounting & Economics*, 1-17. doi:10.1080/16081625.2016.1170100

Smith, A. (1776). *Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan & T. Cadell.

Smith, M. P. (1996, March). Shareholder Activism by Institutional Investors: Evidence from CalPERS. *The Journal of Finance, 51*, pp. 227-252. Retrieved from www.jstor.org/stable/2329308

Social Investment Forum Foundation. (2010). *Report on Socially Responsible Investing Trends in the United States.* Retrieved from http://www.ussif.org/files/Publications/10_Trends_Exec_Summary.pdf

Stern, N. (2007). *The economics of climate change: The Stern review.* Cambridge: Cambridge University Press.

Stock, J. H., & Watson, M. W. (2012). *Introduction to econometrics* (3 ed.). Harlow: Pearson.

Svenning, C. (2003). *Metodboken* (Vol. 5). Eslöv: Lorentz Förlag.

Söderholm, P., & Sundqvist, T. (2003). Pricing environmental externalities in the power sector: ethical limits and implications for social choice. *Ecological Economics, 46*, pp. 333-350. doi:10.1016/S0921-8009(03)00185-X

Tideman, T. N., & Plassmann, F. (2010). Pricing externalities. *European Journal of Political Economy, 26*, pp. 176-184. doi:10.1016/j.ejpoleco.2009.12.006

Tirole, J. (2001, January). Corporate Governance. *Econometrica, vol 69, No 1*, pp. 1-35.

Tobin, J. (1969). A General Equilibrium Approach To Monetary Theory. *Journal of Money, Credit and Banking,, 1*, 15-29. Retrieved from http://www.jstor.org/stable/1991374

United Nations Principles Responsible Investments. (2016, January). *Fiduciary duty in the 21st century: US Roadmap.* Retrieved from https://www.unpri.org/download_report/24188

United Nations Principles Responsible Investments. (n.d). Retrieved from About: What is responsible investment?: https://www.unpri.org/about/what-is-responsible-investment

Williams, R. (2006). Review of Regression Models for CategoricalDependent Variables Using Stata. *The Stata Journal, 6*(2), 273-278. Retrieved from http://www.stata-press.com/books/long2-review.pdf

Williamson, O. E. (1991, June). Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly, vol 36*, pp. 269-296.

9. Appendix

9.1 Correlation matrix

Correlation matrix

	ROA	Tobin's Q	LnAssets	D/E	Cash/Assets	M/B	Sales Growth	R&D/Assets	CapEX/assets
ROA	1.0000								
Tobin's Q	0.4753	1.0000							
LnAssets	-0.0275	-0.2728	1.0000						
D/E	-0.1232	-0.1228	0.1593	1.0000					
Cash/Assets	0.1723	0.3258	-0.2024	-0.1416	1.0000				
M/B	0.0697	0.1117	-0.1146	0.2071	0.0242	1.0000			
Sales Growth	0.1897	0.1022	-0.0476	-0.0124	0.0015	0.0258	1.0000		
R&D/Assets	0.1439	0.4066	-0.1184	-0.0875	0.2415	0.0438	0.0102	1.0000	
CapEX/assets	-0.0097	0.1322	-0.0916	-0.1484	0.0040	0.0477	-0.0062	-0.0969	1.0000

9.2 Further specification Model 1

As introduced in section 4.3.1 the baseline, and most pedagogic explanation of the model, follows the specification:

$$Yi* = Xi\beta i + Zi\beta + ui\ , \qquad Where\ i = 1\ldots, n\ and\ ui: N(0,1)$$

$$Y = 0\ if\ -\infty < Y* < k1$$
$$Y = 1\ if\ k1 < Y* < k2$$
$$Y = 2\ if\ k2 < Y* < k3$$
$$Y = 3\ if\ k3 < Y* < +\infty$$

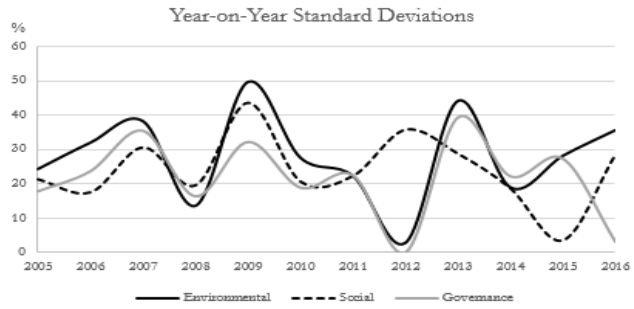
Where Y represents the ordinal variable of Objective result. The Log-likelihood function is constructed as follows. (Daykin & Moffatt, 2002) We set P*i*(Y) as the probability that engagement *i* reaches the result of Y. The probability of this scenario can be defined as:

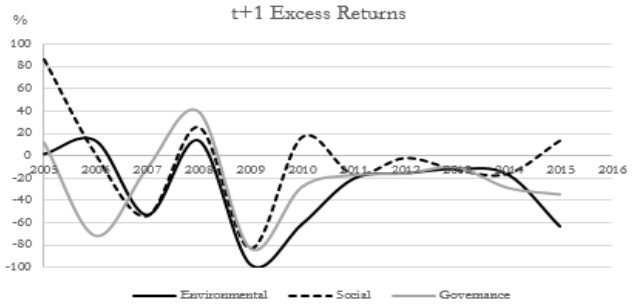
$$Pi(Y) = P(kY - 1 < Yi* < kY) = \Phi(kY - Xi\beta i) - \Phi(kY - 1 - Xi\beta i)\ Where\ Y = 0,1,2,3$$

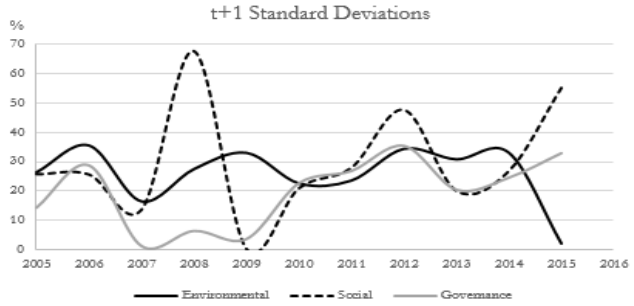
Given a sample of Yi, Xi, i=1...n the log likelihood function for the estimation of Xi particularly becomes:

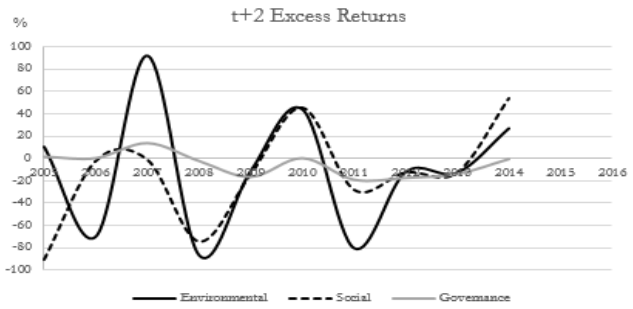
$$LogL = \Sigma ln[Pi(Yi)] = \Sigma ln[\Phi(kYi - Xi\beta i) - \Phi(kY - 1 - Xi\beta i)]$$

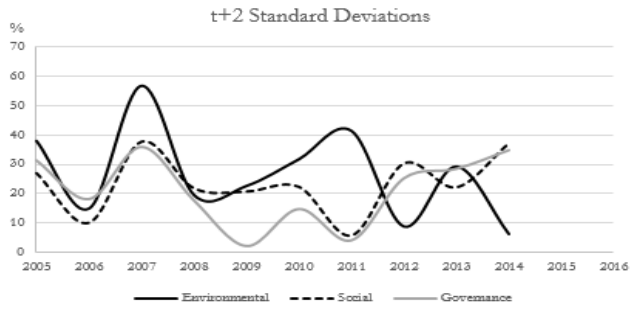
9.3 Graphical display of excess returns from Model 3



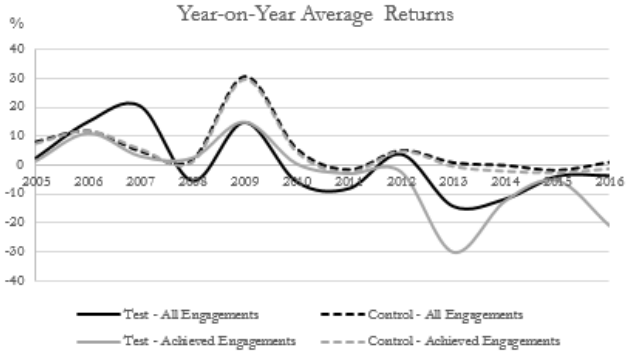


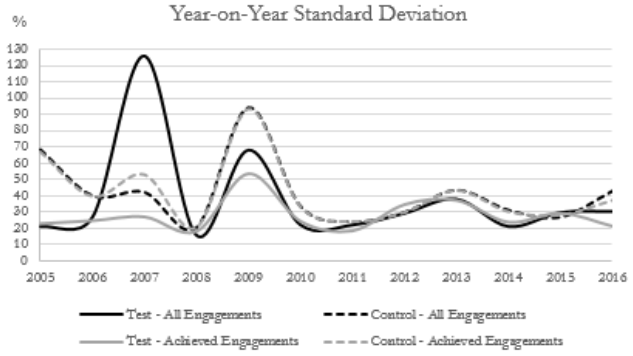


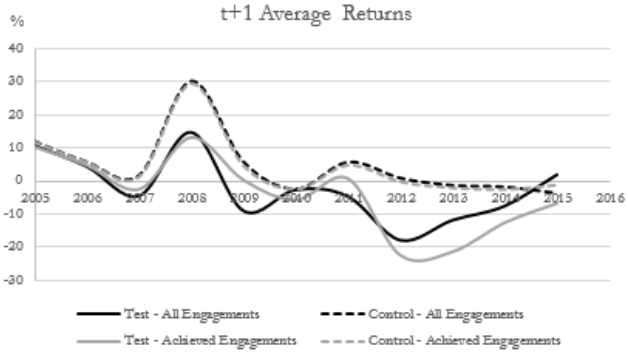


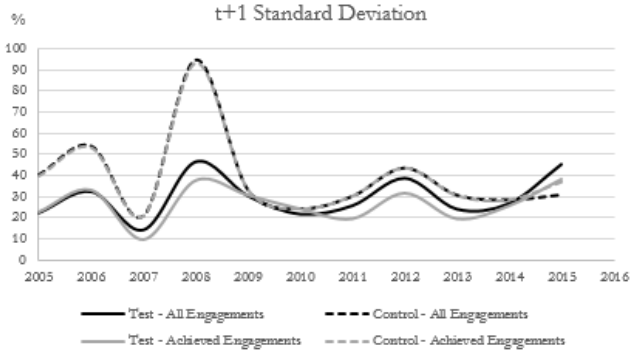


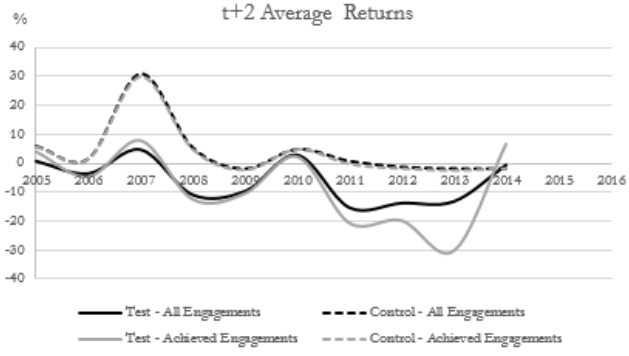
9.4 Graphical display of the comparison between All and Achieved engagements

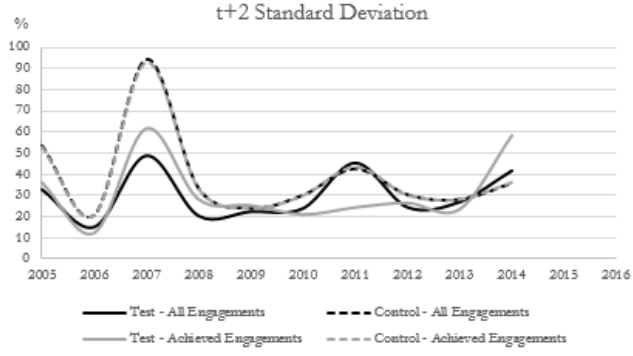












9.5 UNPRI definitions

Environmental

The environmental aspects look at a company's energy use, waste, pollution, natural resource conservation and animal treatment. It also includes evaluating which environmental risks might affect a company's income and how the company is managing those risks.

- **Climate change -** Refers to any significant change in the measures of climate lasting for an extended period of time. According to United States Environmental Protection Agency *"climate change includes major changes in temperature, precipitation, or wind patterns, among other effects, that occur over several decades or longer".*
- **Greenhouse Gas Emissions (GHG) -** Simply defined as gases that traps heat in the atmosphere.
- **Resource depletion (including water) -** Consumption of natural resources faster than it has the possibility to replenish. Commonly refers to areas such as; farming, fishing and water usage.
- **Waste and Pollution -** The definition of waste in accordance to the Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and Their Disposal (1989) *"'Wastes' are substance or objects, which are disposed of or are intended to be disposed of or are required to be disposed of by the provisions of national law".*
- **Deforestation -** Comes in different shapes, including fires, clear-cutting for agriculture, ranching and development, degradation due to climate change and unsustainable logging for timber.

Social

Investigating the company's business relationships. Are they collaborating with suppliers that share the same values? Are they donating a percentage of its profits to the community or perform any kind of volunteer work? Do these companies' working conditions show a high regard for its employees' health and safety?

- **Working conditions (including slavery and child labor) -** The levels of disclosure of human rights policies and processes.
- **Local communities (including indigenous communities) -** According to United Nations, the most fruitful approach is to identify, rather than define indigenous peoples and communities. This is based on the fundamental criterion of self-identification as underlined in a number of human rights documents.
- **Health & safety -** Some companies are facing extremely high financial and reputational risks related to health and safety.
- **Employee relations and diversity -** Can be defined as "*planning and implementing organizational systems and practices to manage people so that the potential advantages of diversity are maximized while its potential disadvantages are minimized*"- Taylor Cox (1993).

Governance

Regarding governance, investors would like to know that a company uses accurate and transparent accounting methods and that shareholders are allowed to vote on important issues. They prefer companies that avoids conflicts of interest in their choice of board members. Lastly, they divest in companies that engages in illegal behavior or use political contributions to obtain favorable treatments.

- **Executive pay -** Remuneration packages should be structured in such way as to optimize financial results and promote sustainable behavior without generating or exacerbating systemic risks that might undermine investors' long-term interests.
- **Bribery and corruption -** For companies, corruption impedes economic growth, distorts competition and represents serious legal and reputational risks. Corruption increases the cost of doing business, and simultaneously raises uncertainty over expected returns for investors.
- **Political lobbying and donations -** A healthy system of corporate governance according to the International Corporate Governance Network is to ensure that the companies make proper use of power that is entrusted to them by their shareholders, especially in terms that includes a company's involvement in seeking influence in the political process.
- **Board diversity and structure -** Director tenure on the boards runs a risk of eroding individual director's independence and objectivity over time. The presence of a significant number of board members with lengthy tenure represents a potential red flag in terms of entrenchment and intransigence.
- **Tax strategy -** Companies are encouraged to disclose information, addressing tax policy, governance and performance to the highest degree possible and with an increasing quality over time.

9.6 Case studies

These tables present three companies, X, Y and Z that have been engaged by Schroders at various points in time. They are included to further deepen the understanding of the engagements and to provide examples of the terminology used in the text.

Company	Engagement date	Objective type	Suggestion for change	Result	E/S/G	Topic	Region	Sector
X	12 Oct 2015	Objective	Asked the company for numerical targets for the number of ships fitted with exhaust gas cleaning technology.	Achieved - The company's 2015 CSR report shows that now 41% of its fleet has been fitted with exhaust gas cleaning technology.	E	Climate Change	North America	Consumer Discretionary
	12 Oct 2015	Objective	Asked the company to start publishing its sustainability report at the same time as its annual report to address the reporting lag.	Achieved - The company has now published CSR reports for fiscal year 2014 and 2015. With the two reports being published in the last 12 months, reporting gap that has existed since 2012 has been closed.	G	Transparency & Disclosure	North America	Consumer Discretionary
	12 Oct 2015	Objective	Asked the company for a 100% compliance target for code of conduct for suppliers.	Almost - As per the company's 2015 CSR report, it has developed a questionnaire to better monitor suppliers' compliance with the company's code of conduct.	S	Supply chain management	North America	Consumer Discretionary
	12 Oct 2015	Objective	Asked the company for health & safety (H&S) data for staff and customers as well as a clear commitment to zero accidents.	Some change - The company's 2015 Corporate Social responsibility report (CSR) now provides more narrative on customer safety and aims for zero accidents. It is not clear whether the injury and accident rates now also include customers as well as staff. This reflects som change on previous reporting.	S	Health & Safety	North America	Consumer Discretionary
	12 Oct 2015	Objective	Asked the company for an absolute target rather than an intensity target, or at a minimum a more stretching CO2 intensity target.	Some change - The company's 2015 CSR report indicates no change to targets, while still working on an intensity target. However, it does now disclose reductions in absolute levels of greenhouse gasses (GHG) from 2011 to 2015.	E	Climate Change	North America	Consumer Discretionary
	12 Oct 2015	Objective	Asked the company for a target concerning employee diversity.	Some cange - The company's 2015 CSR report provides more information on its diversity programmes but does not have any quantifiable metrics as yet.	S	Human Capital Management	North America	Consumer Discretionary
	08 Dec 2014	Objective	Pushed for greater level ROIC targets as a way of sending a message to the industry and other investors alongside greater disclosure. of set targets.	No Change	G	Remuneration	EMEA	Consumer Discretionary
	08 Dec 2014	Discussion	Discussed need to prove strategic focus and improve the economics of the business to increase earnings and ROIC.	No Further Change Required	G	Corporate Strategy	EMEA	Consumer Discretionary
	01 Jan 2011	Objective	Asked the company for disclosure of absolute CO2 figures, as well as setting an absolute target instead of relative reduction target.	Almost	E	Climate Change	EMEA	Consumer Discretionary

Company	Engagement date	Objective type	Suggestion for change	Result	E/S/G	Topic	Region	Sector
Y	29 Oct 2015	Objective	Asked the company to disclose health & safety data	Achieved - Starting 2016, the company disclosed key health & safety performance data, such as the lost time injury frequency and the total recordable case frequency rates.	S	Health & Safety	Pacific	Industrials
	29 Oct 2015	Objective	Asked the company to provide data on number of ethics-related reports from employees and set up an ethics hotline.	Almost - Starting in 2016, the company now discloses a series of human capital data. This includes discrimination complaints, the percentage of voluntary staff turnover, percentage of new hires recruited locally, and details about its whistle-blowing policy	G	Business Integrity	Pacific	Industrials
	01 Jul 2011	Objective	Asked the company to improve carbon risk disclosure	Achieved	E	Climate Change	Pacific	Industrials
	01 Jul 2011	Objective	Asked the company for greater disclosure on human capital management	No Change	S	Human Capital Management	Pacific	Industrials
Z	12 Nov 2015	Objective	Asked the company to provide more detailed breakdown of hotline reports (e.g. by region or issue)	Almost - The company now provides a breakdown of whistleblowing concerns by issue (e.g. business integrity, personell and safety/sustainability). We would welcome further issue and/or regional breakdowns.	G	Business Integrity	Pacific	Materials
	12 Nov 2015	Discussion	Discussed safety costs/investments in light of cost-cutting	No Further Change Required	S	Health & Safety	Pacific	Materials
	12 Nov 2015	Objective	Should engage London Metal Exchange to introduce sustainable category of aluminium.	No Change	E	Environmental Policy/Strategy	Pacific	Materials