STOCKHOLM SCHOOL OF ECONOMICS
Department of Accounting
Master Thesis in Accounting and Financial Management
Spring 2016

CSR disclosures and analyst forecast accuracy: does quality matter?

An empirical study of whether CSR reporting matters to market participants utilising a novel approach to measure reporting quality

Sybren Hent Sjakko Bruin[1]   Daniel Schaffer[2]

Abstract

We investigate whether the quality of Corporate Social Responsibility ("CSR") report narratives affects analyst forecasts, using a sample of reports issued by companies listed in the Nordic countries, the Netherlands and Germany between 2012 and 2014. To quantify the quality of CSR narratives, we draw upon advances in Natural Language Processing techniques to construct a score of CSR reporting quality arguing that better reports are easier to read, longer, contain more forward-looking information and are not overly optimistic in tone. The findings presented in this study contribute to current research in two ways: first, we provide indicative evidence that the quality of CSR report narratives matters to market participants. This suggests that further guidance by standard setters for companies on how to construct their CSR reports would be welcomed by investors. Additionally, the study contributes to a developing body of research employing computerised analysis of corporate disclosure narratives.

Keywords Corporate Social Responsibility reporting, Natural Language Processing, Narrative research, Forecast accuracy

Tutor Henrik Nilsson

Date 2016-05-16

[1] 40719@student.hhs.se
[2] 40709@student.hhs.se

Acknowledgements

We would like to extent our gratitude to our tutor, Henrik Nilsson, for devoting his valuable time to guide us through the process of performing this study. Additionally, we thank Florian Eugster for his suggestions, which helped us find our way early on in the process. We also would like to extent our gratitude to the researchers connected to the Corporate Financial Information Environment ("CFIE") project for providing feedback on how we applied the linguistic analysis tool, Wmatrix, developed under CFIE's umbrella. Finally, we thank Per-Olov Edlund for being available to us when we had questions related to statistical analysis.

Table of Contents

1 Introduction

1.1 *The rise of Corporate Social Responsibility reporting*

Corporate Social Responsibility[1] ("CSR") reporting has its roots in political and societal movements demanding greater insight in the control that multinational companies have over vital resources, and how business operations affect deteriorating environments (Fifka 2013). At the turn of the century, around 50% of Global Fortune 250 companies reported on CSR (Kolk 2003). Recognising that reporting methods were extremely fragmented, the Global Reporting Initiative ("GRI") initiated the development of a global standard on social and environmental reporting (Global Reporting Initiative 2015a). In doing so GRI has helped to establish CSR reporting as common place, with 95% of the Global Fortune 250 companies now reporting on their corporate responsibility (KPMG 2015). Yet, reporting on CSR performance remains largely voluntary, with few countries mandating sustainability reports according to specific frameworks (ESG Analytics 2015). Additionally, several competing frameworks to GRI, such as the UN Global Compact and those issued by the American Sustainability Accounting Standards Board, remain in use. In all, it appears that CSR reporting has become an integral part of the corporate disclosure package, but that reporting styles remain rather fragmented. This raises the question whether CSR reports, of which the quality is likely to differ substantially between report(er)s, contain incremental and relevant information about future firm performance and as such may help to shape investment decisions (Milne & Chan 1999).

One group of stakeholders benefits directly and measurably from the publication of corporate information: financial (sell-side) analysts. The principal job of analysts is to provide forecasts of earnings and share prices to market participants at a cost. Forecast performance is likely to matter to analysts (Mikhail, Walther & Willis 1999), not least because financial analysts' forecast accuracy is related to reputation, and by extension, pay (Stickel 1992). Hence, in an attempt to provide the most accurate forecasts, analysts utilise all types of corporate information sources, including those that are non-financial (Previts et al. 1994). Although only ten years ago equity analysts expressed scepticism regarding the usefulness of sustainability-related non-financial information (Global Institute for Partnership and Governance 2005), their view has widened considerably in recent years (Global Reporting Initiative & Accounting for Sustainability 2012). Therefore, we expect that if CSR disclosures provide useful information to market participants, this is likely to be measurable through the accuracy of analyst forecasts.

[1] We define Corporate Social Responsibility as encompassing all non-financial matters that are affected by business operations. In that sense, it is similar to the "Sustainability", "ESG" and "Stakeholder" movements that call upon business to consider its non-financial responsibilities.

Furthermore, because of their role as information intermediaries, sell-side analysts are likely to shape the opinions of a large number of market actors (Schipper 1991).

Prior research finds ambiguous results regarding the usefulness of CSR reporting. Critics claim that companies make "symbolic use of [CSR reporting] practices" and that CSR disclosure is used primarily as a social legitimation tool rather than reflecting actual CSR performance (Patten 2005, Aras & Crowther 2009). Yet, recent studies support the hypothesis that CSR narratives are useful. Clarkson et al. (2013) claim that CSR reports are incrementally informative and have a signalling effect for firm performance. Additional evidence has been found by other researchers, who contend that non-financial disclosure contributes to more accurate valuations, lower cost of equity and is associated with higher forecast accuracy (Schadewitz & Niskala 2010, Dhaliwal et al. 2011, Dhaliwal et al. 2012).

1.2 The evaluation of CSR reporting in the present study

Several researchers have studied the content of CSR reports (Roca & Searcy 2012, Michelon, Pilonato & Ricceri 2015). However, to our knowledge no study has been published that considers whether quality differences in CSR disclosure affect the forecasts of analysts, while the largely unregulated nature of non-financial disclosures implies that their content differs greatly (Roca & Searcy 2012). Our study aims to bridge this gap and additionally responds to a call for research into the quality of narrative disclosures by utilising advances in Natural Language Processing ("NLP") tools (Beattie & Davison 2013). This type of computerised analysis enables researchers to evaluate large amounts of text, thereby allowing for increased sample sizes compared to manual content analysis studies (Beattie 2014). To interpret linguistic characteristics, we draw upon textual analysis research of other accounting narratives, such as annual reports or earnings press releases (Li 2008, Schleicher & Walker 2010, Li 2010, Lehavy, Li & Merkley 2011, Davis, Piger & Sedor 2012). Based on their findings, we hypothesize that CSR reports of higher textual quality allow analysts to make more accurate earnings forecasts.

We collect all CSR reports published in the period between 2012 and 2014 by public firms listed in the Nordic countries, Germany and the Netherlands that are followed by analysts. Subsequently, we employ an NLP tool to quantify the textual characteristics of these reports and based on previous literature develop a measure of reporting quality arguing that better reports are easier to read, longer, include more forward-looking information and have a less optimistic tone. Inspired by Lang & Lundholm (1996) and Cooke (1998), we combine these four components into one narrative index by summing a report's within-sample rank of each characteristic. We validate our measure of reporting quality by showing that it is significantly related to a measure of CSR performance as well as to whether a report is independently assured. Finally, we divide the reports into three groups representing low, middle and high quality reports and estimate regression models to determine whether higher quality CSR report narratives allow analysts to make more accurate earnings forecasts (Muslu et al. 2014).

We find indicative evidence that forecast errors are negatively related to the quality of narratives within stand-alone CSR reports. The estimated regression coefficients are negative and increase in size from the group of low quality reports to the one with the highest quality, with the average absolute forecast error as a fraction of the share price decreasing by 2.95%, 4.09% and 4.53% compared to non-CSR reporters. However, only the coefficient of the middle group is significant at the 10% level, while the other groups have p-values above 10% but below 20%. Additionally, we find that the continuous narrative index is negatively but insignificantly related to forecast error. Robustness checks provide similar indicative results. Deconstructing the narrative index into its component parts reveals that especially forward-looking information appears to be a useful aspect of CSR disclosure in improving forecast accuracy.

1.3 Implications of our findings

We contribute to the current research body both by exploring a novel method of evaluating the quality of CSR reporting narratives, as well as by providing indicative evidence that better narratives provide more useful information to an important group of market participants, namely financial analysts. We believe that the weak statistical significance found in this study is likely to be the result of the comparably small sample size, but also due to the exclusion of almost 50% of CSR reports because they are rejected by the utilised NLP programme. Nevertheless, by validating our disclosure quality measure we believe that the methodological approach is sound, which, combined with our indicative findings on the relationship between forecast accuracy and CSR reporting quality, should be considered an incentive to continue research in this area. In so doing, we believe that future research might provide an even more convincing case that improving the quality of CSR reporting is useful to market participants.

This paper is organised as follows: section 2 provides an overview of past research into narrative analysis as well as forecast accuracy and concludes with our hypothesis. This is followed by an introduction into our data sample and method in section 3. Section 4 contains uni- and multivariate analyses and robustness tests. Section 5 discusses the limitations of this study and section 6 provides the conclusion.

2 Literature Review and Hypothesis Development

2.1 Literature Review

2.1.1 Scholarly takes on CSR reporting

The first notable contributions on CSR reporting research took place in the 1970's, while a resurgence of research attention is discernible starting after 2000 (Cho et al. 2015). Most research is either concerned with organisational legitimation or investigating the decision usefulness of CSR reports (Gray, Kouhy & Lavers 1995). The former studies the extent to which the decision to publish a CSR report is influenced by organisational legitimation attempts, i.e. "establishing congruence between the social values associated with or implied by [a company's] activities and the norms of acceptable behaviour in the larger social system" (Dowling & Pfeffer 1975, Neu, Warsame & Pedwell 1998). As such, several studies analyse the organisational motivations for- and gains from CSR reporting (Archel, Husillos & Spence 2011, Moser & Martin 2012, Cho et al. 2015). Decision usefulness studies attempt to determine whether CSR information allows users of corporate information to make more informed decisions. They are typically conducted as ranking or information effect studies, but have produced ambiguous results during the first period of CSR research (Gray, Kouhy & Lavers 1995).[2] However, recent studies show that CSR information appears to be useful for the decisions made by market participants (Guidry & Patten 2010, Dhaliwal et al. 2011, Dhaliwal et al. 2012, Clarkson et al. 2013). Additionally, Pflugrath, Roebuck & Simnett (2011) show that firms improve the credibility of information contained in their CSR reports by acquiring external assurance, a costly exercise which is unlikely to be undertaken if the information would not be used in decision making.

Some researchers have taken a very critical perspective on CSR reporting, claiming that companies use CSR reporting as an obfuscation or 'greenwashing' technique, whereby firms deliberately disguise or misrepresent their environmental performance to manage reputational risk (Patten 2005, Delmas & Burbano 2011, Michelon, Pilonato & Ricceri 2015). Nevertheless, several studies underscore the 'truthfulness' of CSR reports, finding links between sustainability reporting and sustainability performance (Clarkson et al. 2008).

[2] Ranking studies generally ask decision makers (e.g. analysts) to 'rank' information sources depending on their perceived usefulness, while information effect studies aim to directly assess the effect of released information by estimating how it influences observable variables such as analyst forecasts or share price.

2.1.2 Assessing the quality of CSR disclosure narratives

As described in section 1.1, CSR reporting has become an integral part of the narrative disclosure package utilised by firms. To find theoretical guidance on how to analyse CSR disclosures, we therefore turn towards the corporate disclosure literature. One venue of disclosure quality research is the usefulness of reported financials, an area which receives extensive consideration through the Earnings Management literature (see Healy & Wahlen (1999) for an overview). However, there is debate among scholars whether "financial information is losing its relevance" (Francis & Schipper 1999). As such, a number of studies extend the analysis of corporate disclosures beyond the reported financials to exploit the full range of information sources available.

When attempting to evaluate the complete information environment, including such disclosures as press releases and analyst calls, researchers have frequently employed disclosure scores developed by third party institutions in order to assess the quality of non-financial information (Lang & Lundholm 1996, Hope 2003a). Such scores also exist in the field of CSR reporting, notably regarding CSR information published by US companies (KLD ratings) and recently utilised in a study by Gao et al. (2015), of CSR disclosures by organisations listed in the Netherlands. However, third party ratings are criticised for their subjectivity and low reproducibility (Beattie, McInnes & Fearnley 2004), and are furthermore limited to the sample chosen by the rater.

Theory underlying the research on narrative disclosures

Acknowledging the disadvantages of third-party disclosure scores, some researchers choose to evaluate narrative disclosures directly (Jones & Shoemaker 1994). In her literature review on research in the field of accounting narratives, Beattie (2014) distinguishes two major research strands, namely disclosure research and narrative research. Disclosure research is centred on the question of 'what' information is disclosed. It analyses the financial and qualitative information provided through corporate disclosures and scrutinises whether this data reduces information asymmetries (Healy & Palepu 2001). The second stream, narrative research, focuses on the language in qualitative disclosures and analyses how narratives create meaning for 'human actors'. Narrative sections allow management to deliberately steer the tone and content of the presented information, since there are few guiding principles on how to write them (Merkl-Davies & Brennan 2007). In their comprehensive literature review on the motivation of narrative disclosures, Merkl-Davies & Brennan (2007) distinguish between impression management, with the intention of obfuscating the reader, and the provision of incremental, i.e. useful and reliable, information. Sustainability reporting is exceptionally prone to the application of impression management due to the difficulty of quantifying environmental performance (Neu, Warsame & Pedwell 1998, Roca & Searcy 2012). Neu, Warsame & Pedwell (1998) study the role of discretionary environmental reporting and find that impression management might be prevalent, but acknowledge that their results are inconclusive.

Methodologies to study narrative disclosures

Beattie (2014) identifies three methodological approaches to study narrative disclosures: discourse, content and linguistic analysis.

Discourse analysis has been used in different scientific disciplines and refers to the application of language and the creation of meaning taking into account the specific societal background of actors (Alvesson & Karreman 2000). Content or thematic studies draw on the different topics covered in corporate disclosures and are traditionally performed by scoring reports based on pre-defined disclosure indices.[3] Disclosure index studies were first used in the field of accounting in the 1970s, but came to prominence after Botosan (1997) published her seminal paper reporting a negative relation between improved corporate disclosures and cost of equity capital. With regards to CSR reporting, Patten (2002) finds a significant relationship between environmental disclosure and environmental performance. Furthermore, Clarkson et al. (2008) develop a disclosure index score using GRI guidelines as guidance, and find a positive relation between disclosure and environmental performance for their 191 sample companies operating in polluting industries in the USA. More holistic studies go beyond simply scoring the presence of certain topics, but additionally measure richness as a product of width and depth by analysing the context in which topics are presented (Beretta & Bozzolan 2008). Similarly, Beattie, McInnes & Fearnley (2004) develop a four dimensional model for holistic content analysis using a computer aided analysis program. Common methodological problems with content analyses are limited reproducibility and divergent research approaches due to the need for subjective assessments regarding quality (Jones & Shoemaker 1994, Botosan 1997, Beattie, McInnes & Fearnley 2004). Furthermore, as content analysis requires the researcher to manually screen each report, sample sizes are generally limited.

Linguistic or textual analyses scrutinise specific textual features of corporate narratives, while not necessarily examining the delivered content of reports. Li (2010) categorises common

[3] In these studies, researchers define a list of relevant topics and score the report based on the information provided on these topics. Scoring could be binary (0=not covered; 1=covered) or categorical in nature, where an integer score is awarded based on how much (if any) qualitative or quantitative information is provided.

research approaches into readability, disclosure level (i.e. the amount of disclosure) and disclosure tone. Compared to content analysis, these methods allow for larger sample sizes since data can processed automatically using NLP techniques, which additionally increases objectivity and reproducibility (Beattie 2014).

Linguistic Analysis: Readability

Readability (sometimes referred to as transparency) studies attempt to determine how the legibility, and to some extent understandability, of a text affects decisions made by the reader.[4] Indicators of readability frequently employed are the Fog- or Flesch-Index.[5] Courtis (1998) uses the Flesch-Index to analyse whether firms deliberately employ difficult language to obfuscate bad news. However, the sample of 120 companies listed in Hong Kong does not confirm his hypothesis. In contrast, Li (2008) finds that firms with less readable annual reports publish lower earnings while firms with better readable annual reports have positive persistent earnings. Furthermore, Lehavy, Li & Merkley (2011) document that readability of 10-K filings affects forecast accuracy, analyst following and the informativeness of analyst reports. Work on the readability of CSR reports was conducted by Bakar, Sheikh & Ameer (2011). Their analysis of 333 Malaysian companies shows a positive association between performance and readability, supporting the obfuscation hypothesis.

Linguistic Analysis: Disclosure level

Researchers have furthermore investigated the relationship between disclosure quality and quantity. Many studies apply different proxies such as word-, sentence- or page count for disclosure quantity (Hackston & Milne 1996). While Beretta & Bozzolan (2008) claim that disclosure quality is not a function of total disclosure, Hooks & van Staden (2011) find a significant relationship between disclosure quality and disclosure quantity for environmental reports. Similarly, Kothari, Li & Short (2009) discuss the application of word counts to compare texts and to compare keyword occurrences across texts by contrasting the fraction (number of keywords over total number of words) between written narratives.

[4] Jones & Shoemaker (1994) note that understandability is also shaped by the reader's cognitive abilities.

[5] The Fog Index measures readability based on the number of syllables per word and the number of words per sentence, resulting in the number of years of education necessary to understand a text, meaning the lower the score the better readable the text (Loughran, & McDonald 2011). The Flesch Index uses the number of syllables per 100 words and the sentence length to compute a score from 0 to 100, with the higher number indicating better readability (Sydserff & Weetman 1999).

Linguistic Analysis: Disclosure tone

Disclosure tone can be analysed by counting the number of occurrences of specific words in a text, thereby attempting to identify specific language patterns that reveal information about the writer's actual intentions or perspective. Counting the number of pre-defined keywords (a 'bag of words') of a certain category (e.g. positive, negative or forward-looking vocabulary) can be done manually or with NLP software. Studies usually apply a standardised bag of words, such as the Harvard Psychology Dictionary (Kothari, Li & Short 2009). However, Loughran & McDonald (2011) find that this list is not suitable for corporate texts, since words might be misclassified due to different contextual meanings.

Cho, Roberts & Patten (2010) report that firms with worse environmental performance apply a more optimistic and less certain tone in their reports.[6] Holder-Webb et al. (2009) note a "generally self-laudatory tone in the content of CSR disclosures for their sample firms [in the USA]." However, Arena, Bozzolan & Michelon (2015) argue that more optimistic disclosures could in fact be a sign of better future performance. Besides optimism and certainty, studies on narrative tone find an association between forward looking statements in the Management Discussion & Analysis ("MD&A") part of annual reports and forecast accuracy or firm performance (Hussainey, Schleicher & Walker 2003). Referring to this study, Beretta & Bozzolan (2008) find evidence that forward looking information is associated with the extent of forecast revisions by analysts during the forecasting year. These results were extended by a statistical study by Li (2010), finding a statistical relationship between forward-looking statements and future earnings and liquidity.[7] Other studies develop an individual bag of words by parsing a sample of narratives for frequently used words and adding synonyms (Hussainey, Schleicher & Walker 2003).

2.1.3 Measuring the decision usefulness of disclosures through analyst forecasts

We have identified that the usefulness of additional corporate narrative disclosures contained in CSR reports can be analysed using various methodologies. The question remains, however, which users would benefit from such an analysis. Since CSR reporting has been adopted by nearly all of the world's largest firms (KPMG 2015), a starting point is to identify the most important users of these firm's (financial) disclosure.

[6] Optimism being defined as emphasising positive news/events while concealing negative news/events.

[7] The statistical approach (as against to the dictionary approach) to narrative tone assessment "relies on statistical techniques to infer the content of text and classify documents based on statistical inference" (Li 2010).

Generally, researchers assume the primary users of corporate disclosures to be financial analysts, as "given their importance as intermediaries who receive and process financial information for investors, it makes sense to view analysts as representative of the group to whom financial information is and should be addressed" (Schipper 1991). Analysts act as information intermediaries and sell investment advice in the form of buy/hold/sell recommendations as well as profit forecasts for analysed firms (Fogarty & Rogers 2005). Furthermore, in an attempt to provide the most accurate forecasts, analysts utilise all types corporate information sources, including those that contain non-financial information (Previts et al. 1994). Guided by Ramnath, Rock & Shane (2008), we therefore analyse several usefulness studies that evaluate how (non-)financial information affects decisions made by analysts.

The usefulness of better disclosure to analysts: annual reports

Given the nature of their activities, better information disclosure is only relevant for analysts if it affects the value or profit of the company on which they are writing a report. Intuitively, one would argue that this information is primarily quantitative. However, Holland (1998) notes that qualitative disclosure is used by firms to help investors understand their business and make assessments about future performance. Furthermore, an increasing number of information sources (e.g. the Internet) and the expanding MD&A sections of annual reports has led researchers to question whether "Financial Statements Lost Their Relevance" (Francis & Schipper 1999). This realisation has spurred research into the usefulness of additional disclosure besides the financial information content of annual reports.

Following Lang & Lundholm (1996), many researchers have assessed the relevance of additional information to analysts by estimating how it affects the accuracy of their profit forecasts. Lang & Lundholm (1996) find a positive relation between disclosure scores assigned by the US Financial Analyst Federation and the accuracy of analyst forecasts. Directly testing the usefulness of non-financial information, Barron, Kile & O'Keefe (1999) report that forecast accuracy increases and dispersion falls for companies with higher SEC scores of MD&A disclosure. Extending the analysis to include non-US companies using CIFAR scores, Hope (2003a) documents a similar positive relationship between disclosure scores and forecast accuracy.[8] He also notes that strong enforcement is associated with higher forecast accuracy. In a more recent Italian setting, Beretta & Bozzolan (2008) assess the usefulness of forward-

[8] The Center for International Financial Analysis and Research ("CIFAR") scored the corporate disclosures of companies operating in a multitude of countries. It has a similar function as the USA Financial Analyst Federation ("FAF"), whose reports have also been used extensively in disclosure research.

looking disclosure sentences by employing a disclosure index that includes both qualitative and quantitative aspects, and find a statistical relationship between their measure and analyst dispersion and accuracy.

It must be noted that researchers have identified many additional factors that drive analyst forecast accuracy besides disclosure quality.[9] Performing a meta-analysis study of several analyst forecast accuracy studies, Garcia-Meca & Sanchez-Ballesta (2006) show that firm size and forecast horizon are consistently found to be explanatory variables in forecast accuracy studies. Furthermore, Hope (2003b) unveils the relevance of culture, while Lys & Soo (1995) argue that larger analyst following leads to an 'accuracy competition'.

The usefulness of better disclosure to analysts: CSR reporting

As documented in section 1.1, companies might report on CSR to address the concerns and requirements of various stakeholder groups. Especially relevant for the work of analysts, however, is that CSR activities could potentially have a positive effect on a company's financial performance through a number of strategic channels: lower costs as a result of increased efficiency, higher productivity due to increased employee satisfaction or the ability to hire better workers as well as generate additional revenue through increased customer loyalty or acquisition of new customers (Burke & Logsdon 1996).

Burke & Logsdon's (1996) assertions are furthermore confirmed by Orlitzky, Schmidt & Rynes (2003), who find a generally positive relationship in a meta-analysis of quantitative CSR-financial performance relationship studies. Similar to enhanced financial disclosures, CSR disclosures seem to reduce the inherent information asymmetry between managers and investors, with investors rewarding managers that provide more forthcoming disclosures with lower cost of equity (Dhaliwal et al. 2011). In addition, the authors find that firms publishing their first CSR report attract additional analysts and are able to raise more capital than others. Cheng, Ioannou & Serafeim (2014) confirm the findings regarding cost of capital and link them to an increased stakeholder engagement and transparency of CSR reporting firms. Kim, Park & Wier (2012) find evidence that CSR reporting firms are less inclined to engage in earnings management or other kinds of manipulation. With regards to analysts' general perception of CSR reporting, Ioannou & Serafeim (2015) report a shift by financial analysts, from pessimistic recommendations for CSR reporting firms in the 1990's towards optimistic recommendations

[9] For an overview of the financial forecasting literature published in the period between 1992 and 2006, please refer to Ramnath, Rock & Shane (2008).

until 2007, also noting that this shift was driven by “more experienced analysts as well as higher-status brokerage houses”.

Potentially adverse CSR events, on the other hand, can have serious negative financial consequences: BP has paid over $50bn in fines and clean-up costs as a result of the Deepwater Horizon oil spill (Gilbert & Kent 2015). Additionally, CSR investment costs might outweigh the benefits. This assertion is supported by Brammer, Brooks & Pavelin (2006), who find a negative relation between corporate social performance and stock prices of UK companies.

Given the potential (positive or negative) effect of CSR performance on firm financial performance, disclosure of a firm’s CSR activities is likely to be useful information for analysts. In an extensive international study Dhaliwal et al. (2012) indeed document a significant positive relationship between the issuance of a stand-alone CSR report and forecast accuracy. Their study furthermore reveals that opaque financial disclosure and the stakeholder orientation of a country strengthen the relationship between CSR disclosure and forecast accuracy, highlighting that CSR reports complement the existing information environment. In addition, Cormier & Magnan (2014) report that improved environmental and social disclosure (measured using a disclosure index) by Canadian firms decrease the dispersion of analysts’ forecasts, while a working paper by Muslu et al. (2014) suggests that higher quality CSR reports results in lower forecast errors in the USA. Finally, Harjoto & Jo (2015) use American CSR performance data and find that increased CSR intensity is associated with lower analyst forecast dispersion, but that the effect diminishes along increased disclosure quality.

2.1.4 Research gap

In this section, we identified four different approaches to the evaluation of corporate disclosure: employing third party scores, or evaluating narratives directly by means of discourse, content or linguistic analysis. However, especially third-party disclosure scores and content analysis suffer from low reproducibility, subjectivity as well as limitations in sample size. Furthermore, discourse analysis requires repeated interaction with relevant actors. With this in mind, one promising direction of research is linguistic analysis assisted by computer aided NLP techniques, which offers faster and more objective analyses of readability, disclosure level and disclosure tone, i.e. impression management. It also allows for much larger sample sizes (Core 2001, Hussainey, Schleicher & Walker 2003). Following Beattie & Davison’s (2013) call for further research on accounting narratives, including those embodied in social responsibility reports, we see large potential for future research to combine the benefits of large-scale NLP studies and the increasing universe of CSR reports. Lastly, it was shown that financial analysts are the intended audience of a large part of corporate disclosure. As such, we believe that they are among the primary beneficiaries from improved CSR disclosure quality, allowing them to make more accurate profit forecasts. We thus believe there to be room for research that evaluates the usefulness of CSR disclosures to analysts, utilising advances in NLP techniques. This research gap is illustrated in Exhibit 1.

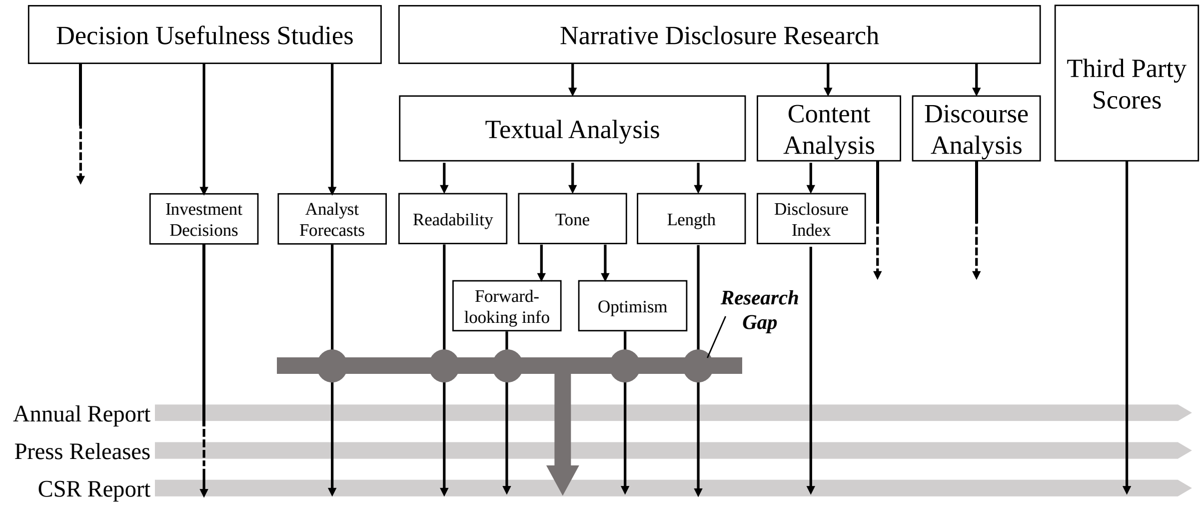


Exhibit 1: Research Gap

2.2 Hypothesis development

Different contemporary studies have already demonstrated the positive signalling effect of CSR disclosure on the accuracy of firm valuation, cost of capital and analyst forecasts. In this study, we will go beyond this knowledge and analyse the different characteristics of stand-alone CSR reports. In so doing, we shed light on the question whether variations in the quality of disclosures contained in CSR reports affect the accuracy of analyst forecasts. Given the prevailing concerns regarding obfuscation or 'greenwashing', we assume that CSR reports entail varying degrees of informativeness. Therefore, we conjecture that the amount of relevant information analysts are able to derive for their earnings forecast differs across reports, and expect that better narrative information in turn allows analysts to make more accurate forecasts. Therefore, the following hypothesis will be the focus of our study:

"Higher quality CSR disclosure narratives allow analysts to make more accurate earnings forecasts."

Based on findings of previous narrative research, we furthermore conjecture that more informative, i.e. better, CSR reports are distinguished by specific positive textual characteristics. If we find that CSR disclosure enables analysts to make more accurate earnings forecasts, we contend that this provides indication that CSR information is useful to market participants. However, we believe that this usefulness is a function of the quality of the narratives in which CSR information is presented.

3 Research Method

3.1 Meta design

We apply a deductive research approach by employing extant literature to guide the development of our theoretical framework and frame our hypothesis, which we subsequently test by performing quantitative analyses (Hyde 2000). In our review of the previous literature, we find that even though CSR reporting has been subject to scientific research for a long time, it has rarely been analysed from a linguistic perspective. Drawing from linguistic theory, we learn that textual characteristics impact how readers perceive the information content of written narratives. Therefore, we test whether this theoretical construct extents to CSR reports by measuring the effect of narrative quality on financial analyst forecast accuracy. We do so based on findings from prior decision usefulness studies, which provide evidence that increased non-financial disclosure allows analysts to make more accurate profit forecasts.

We employ a quantitative research design, which allows us to draw verifiable and falsifiable conclusions about the decision usefulness of CSR reports of different quality. More specifically, the present study is of correlational design, since we use correlational statistics (regression analysis) to make inferences from our observations (Creswell 2013). Accordingly, this study is rooted in post-positivistic epistemology, as we hope that our sample represents the total population and thus allows us to draw generalisable conclusions about the relationship between CSR reporting quality and analyst forecast accuracy (Guba & Lincoln 1994). Following post-positivist epistemology, we use observable evidence and rational considerations to develop our knowledge about the subject and describe the relationship between the objects we study (Creswell 2013).

3.2 Data collection

3.2.1 Collection of forecast information

In order to test our hypothesis, we start our data collection procedure by obtaining a list of all companies traded on stock exchanges in the Nordic countries (excluding Iceland) plus Germany and the Netherlands from Thomson Reuters Datastream. Although the focus of this study originally laid only on the Nordic countries, the resulting sample was deemed too small to draw valid inferences. Therefore, Germany and the Netherlands were added because they are considered to be relatively similar to the Nordic countries.[10] We retrieve all forecast information that concerns these companies for the years 2013 to 2015 from the Institutional Brokerage System Database ("I/B/E/S"), which we accessed through Wharton Research Data Services ("wrds"), and excluded all companies without such information. After deducting observations for which we could not obtain the necessary control variables, we identified 3 048 firm-years of data from 1 065 unique companies, which forms our potential pool of observations.

3.2.2 Collection and classification of CSR information

Upon establishing the potential pool of observations, we employ various sources to identify whether a company issued a stand-alone CSR report for the years 2012, 2013 and 2014.[11] Firstly, we use the GRI Report List published by the Global Reporting Initiative which contains "all sustainability and integrated reports that the GRI is aware of" (Global Reporting Initiative 2015b). As we find that this list suffers from various limitations, we supplement it by manually visiting the website of each company to classify its yearly reporting. Any downloadable report detailing the communication of CSR activities during a specific time frame is classified as a stand-alone CSR report. Note that all firm-year observations classified as having a CSR report also publish financial information in an annual report ("AR"). We include 'Communication on Progress' ("COP") reports, which are required to be published by signatories of the United Nations Global Compact initiative. We include COP reports because the initiative strives for its signatories to produce meaningful CSR reports, which is captured in its conviction that "reporting to stakeholders in a transparent and public manner [is] fundamental for companies committed to sustainability" (UN Global Compact 2015). However, COP reports as well as the mandatory CSR reports required under Danish law are only classified as such when it is deemed that the respective company had made an honest attempt at reporting CSR issues, indicated by publishing at least five pages.[12]

[10] An additional reason for including the Netherlands and Germany is that the researchers have personal affinity with these countries.
[11] For firms where the financial year does not coincide with the calendar year, the year label assigned in this study would be the year in which the respective financial year ended. I.e., a company's report that is labelled as referring to the 2011/2012 financial year would be included as a report referring to the year 2012 in this study. We ensure that all data gathered is in line with the respective financial year-end dates.
[12] Note also that companies might produce CSR reports bi-annually, in which case we consider those reports to be only relevant for the latest reporting period.

To avoid mistakenly assuming that a company does not report sustainability information if it does not publish a separate CSR report, we furthermore screen and classify ARs that contain substantial CSR information as integrated reports ("IR"). Integrated Reporting is a new type of reporting advocated by the International Integrated Reporting Council, which advocates that the reporting of sustainability and financial information is combined into one publication (International Integrated Reporting Council 2014). We assign this classification either when a report is titled 'Integrated' or when at least 10% of the pages in the AR are dedicated to the description of CSR activities (usually contained within a section called 'Sustainability Report').[13] We use this second criterion because we find that several, especially smaller, companies seem to be devoted to publishing significant amounts of CSR related information in their AR without specifically labelling it an IR. We hypothesise that this is primarily due to a lack of awareness of the integrated reporting movement and not because the reports include fundamentally different CSR information than those that are labelled as integrated.

Finally, we determine whether a CSR report or IR is externally assured by checking whether an assurance statement is present in either the report or on the company's website. Firm-year observations for which we have forecasts and control variable data, but in which an AR is published serve as the control group in our regression models.

3.2.3 Description of the pool of potential observations

As illustrated in Exhibit 2 – Panel A, we identified 597 (20%) stand-alone CSR reports and 230 (8%) IR reports, implying that we have 2 221 (72%) firm-years of observations without CSR information. Not surprisingly, the German stock market is the largest and most well covered, supplying 36% of potential observations. Interestingly, the second largest amount of observations comes from Sweden while Norway is third, both ahead of the Netherlands which is larger in both GDP and population. We corroborate these findings and find that Nordic countries indeed have a larger number of listed companies than the Netherlands (Federal Bank of St. Louis 2016). Furthermore, we have the highest number of potential observations from companies operating in the Industrials industry, while we collect only very few data points on firms operating in the Utilities or Telecommunications industries.

The number of yearly observations is nearly identical for 2012 and 2013, while the decrease in 2014 is not caused by a sudden spike in defaults but rather is explained by the fact

[13] Occasionally, firms publish extensive CSR information on their website but do not present it as a separate CSR report, nor is the information downloadable. In most cases, we labelled these observations also "IR" to avoid measurement errors regarding the availability of CSR information.

that the calculation of forecast errors requires reported actuals, which were not published by all companies at the time that the data collection for this study was concluded.[14] We realise that this sampling problem might slightly bias our results because the companies that had not reported their results are possibly smaller companies or companies that had to delay the publication of their financial reports for various reasons.[15] We also note that reporting on CSR information continues to be on the rise, with the share of companies that publish stand-alone CSR reports increasing from 18% in 2012 to 23% in 2014, and integrated reporting rising from 5% to 10%. Again, it must be noted that the numbers for 2014 might be inflated due to data selection issues, but the nominal numbers of both IRs and CSRs have also increased steadily over the 3 years.

Exhibit 2 – Panel A furthermore shows that companies listed in Denmark were most active in CSR reporting, which is to be expected since CSR reporting is mandatory for Danish companies with more than EUR 19 million in assets, EUR 38 million in revenues and more than 250 employees (at the 2009 DKK-EUR exchange rate) (Initiative for Responsible Investment 2015). Based on our sample, these requirements appear to be met for companies at the 25 percentile (untabulated), implying that the majority of Danish companies are required by law to publish CSR reports, or state explicitly why they choose to not do so. Companies listed in Norway or Germany were least likely to publish a stand-alone CSR report. Also notable is the high uptake of integrated reporting in the Netherlands and Finland. Similarly, companies operating in the consumer goods industry publish the most CSR reports, perhaps suggesting that the social legitimation explanation for CSR reporting holds some ground. Within our sample, technology and health care firms were relatively least likely to publish a stand-alone CSR report or an integrated report.

[14] As of April 1st, 2016 when the data collection for this study was finalised.

[15] Since size is found to be significantly inversely related to forecast errors later on in this study, it is likely that published average forecast errors are understated.

Observed distribution of reporting type and derivation of final sample

Panel A: distribution of reporting type of available observations

	Denmark			Finland			Germany			Netherlands			Norway			Sweden			Total						per Year	
Year	AR	IR	CSR	AR	IR	CSR	AR	IR	CSR	AR	IR	CSR	AR	IR	CSR	AR	IR	CSR	AR		IR		CSR		**all**	*in %*
2012	31	4	30	49	14	30	337	12	54	55	15	19	147	4	17	216	10	43	835	*77%*	59	*5%*	193	*18%*	**1 087**	*36%*
2013	26	6	29	47	15	29	329	15	52	54	16	19	126	10	21	199	16	47	781	*74%*	78	*7%*	197	*19%*	**1 056**	*35%*
2014	23	4	31	42	19	29	230	20	56	43	23	15	112	10	24	155	17	52	605	*67%*	93	*10%*	207	*23%*	**905**	*30%*
Industry																									**per Industry**	
Basic Materials	3		2	5	6	15	37	14	16	7	6	6	11	6	6	43	5	14	106	*52%*	37	*18%*	59	*29%*	**202**	*7%*
Consumer Goods	9		15	20	7	10	85	4	32	14	4	4	31	3	13	35	11	21	194	*61%*	29	*9%*	95	*30%*	**318**	*10%*
Consumer Services	10			11	8	9	109	3	15	10	6	12	5	2	4	45	5	27	190	*68%*	24	*9%*	67	*24%*	**281**	*9%*
Financials	19	2	11	13	6	9	132	4	34	28	12	9	89	3	5	72	7	21	353	*74%*	34	*7%*	89	*19%*	**476**	*16%*
Health Care	20	7	20	9		3	93		4	12			22			109	1	3	265	*87%*	8	*3%*	30	*10%*	**303**	*10%*
Industrials	15		39	40	17	33	203	11	38	46	18	12	84	4	11	128	13	38	516	*69%*	63	*8%*	171	*23%*	**750**	*25%*
Oil & Gas		3			3		34	2	7	2	6	4	110	4	20	23		3	169	*76%*	18	*8%*	34	*15%*	**221**	*7%*
Technology	1		2	38		6	172	6	7	28		6	27		2	96		10	362	*90%*	6	*1%*	33	*8%*	**401**	*13%*
Telecommunication		2	1	2		1	17		3	5	2		3	2	1	11	1	5	38	*68%*	7	*13%*	11	*20%*	**56**	*2%*
Utilities	3				1	2	14	3	6				3			8			28	*70%*	4	*10%*	8	*20%*	**40**	*1%*
total	80	14	90	138	48	88	896	47	162	152	54	53	385	24	62	570	43	142	2 221	*73%*	230	*8%*	597	*20%*	**3 048**	
in %	*43%*	*8%*	*49%*	*50%*	*18%*	*32%*	*81%*	*4%*	*15%*	*59%*	*21%*	*20%*	*82%*	*5%*	*13%*	*75%*	*6%*	*19%*								
per Country			**184**			**274**			**1 105**			**259**			**471**			**755**	**3 048**							
in %			*6%*			*9%*			*36%*			*8%*			*15%*			*25%*								

Panel B: derivation of final sample

	No. of Reports	No. of Companies
Total number of observations	**3 048**	**1 290**
Exclusion of IR's	(230)	(108)
Observations from AR or standalone CSR reports	**2 818**	**1 182**
Exclusion of non-working CSR reports	(296)	(117)
Final sample	**2 522**	**1 065**
Of which:		
CSR	*301*	*163*
AR	*2 221*	*902*

Notes

AR: firm-year observation in which only an Annual Report (containing financial information) is published.
IR: firm-year observation in which one report containing both CSR and financial information is published.
CSR: firm-year observations in which both a separate CSR report and an (annual) report containing financial information are published.

Exhibit 2: Observed distribution of reporting type and derivation of final sample

3.2.4 The final sample

Although our analysis tool (described later) was designed to be able to handle the unstructured nature of narratives included in the (IFRS based) annual reports published by UK companies, unfortunately it still imposes several stringent requirements for a report to be analysed correctly.[16] These stringent requirements implied that almost 50% of the available reports (296 in total) could not be read, limiting the sample of reports for which we obtain characteristics of the narratives included in those reports to a total of 301 firm-year observations. To be able to make valid inferences about whether sustainability narratives allow financial analysts to make more accurate profit forecasts, we therefore exclude the non-working reports from our main regressions. Employing a similar argument, we also exclude companies that publish an IR, since we believe that CSR narratives contained within annual reports are not comparable to those included in stand-alone CSR reports. From the original 3 048 firm-year observations we hence deduct 230 IR reports and 296 CSR reports that did not work, resulting in a final sample containing 2 522 observations.

3.3 Main research variables

3.3.1 Analyst forecast accuracy

We follow Dhaliwal et al. (2012) and define analyst forecast error as the inverse of forecast accuracy. Thus, we compute forecast error as the average of the absolute difference between each analyst's forecast of EPS and realised EPS, scaled by the stock price at the beginning of the year to facilitate comparison across companies:

$$FERROR_{i,t} = \frac{1}{N}\sum_{J=1}^{N}\frac{\left|FC_{i,t,j}-EPS_{i,t}\right|}{P_{i,t-1}} \tag{1}$$

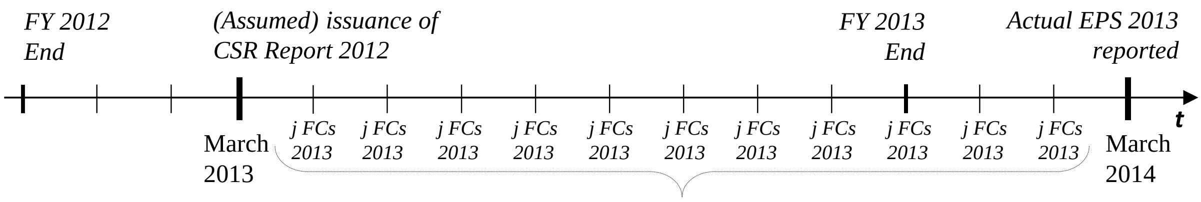
In Equation (1) *i*, *t* and *j* denote company *i*, year *t* and forecast *j*, respectively. *FC* is the forecasted and *EPS* the actual earnings per share, both obtained from the I/B/E/S summary statistics database to ensure consistency.[17] Share price *P* is obtained from Thomson Reuters Datastream, a database created by the same firm that also compiles the I/B/E/S data thus minimising the risks of inconsistencies. We also calculate forecast error according to a different

[16] For example, reports need to contain a content page to be handled properly and need to be written in English.
[17] In some instances I/B/E/S did not provide actuals, while manual inspection revealed that they do exist. To maximise the sample size we therefore employed Datastream to obtain actuals that were not recorded in the I/B/E/S database, ensuring consistency by cross-checking currencies and subjecting the results to a reality check.

method devised by Hope (2003a) but find that using this method does not affect the tenor of the results reported in this study.[18]

Like Dhaliwal et al. (2012), we use all forecasts made during the year to capture the likelihood that the impact of CSR information on EPS is better understood over time.[19] Furthermore, we do so because manual inspection of a subsample revealed that reports need not to be published on the same day as the annual report. However, we believe that the annual report publication date is a good proxy and by including all observations we attempt to average out the measurement error that some of the included forecasts are made without the corresponding CSR information.

Note that we are interested in the analyst forecast error during the year following the publication of a CSR report, as we hypothesise that those reports contain meaningful information for analysts when forecasting the coming year's EPS. Provided in Exhibit 3 is a graphical representation of the process for a hypothetical company *X* in year 2013, assuming that the financial year runs from January to December and that the company publishes its reports in March the year following.



$$FERROR_{X,2013} = \frac{1}{N}\sum_{j=1}^{N} \frac{\left|FC_{X,2013,j} - EPS_{X,2013}\right|}{P_{X,2012}}$$

Exhibit 3: Timeline underlying *FERROR* calculation

3.3.2 CSR report disclosure quality

To evaluate the quality of the narrative disclosures included in stand-alone CSR reports we employ an off-the-shelve software tool named Wmatrix. The tool has been developed by Rayson (2009) as a web based text processing tool, and has been used in language and literature studies (Adolphs 2006). Recently, a team of researchers under the umbrella of the UK Corporate

---

[18] Hope (2003a) calculates forecast error as "the negative of the absolute difference between actual EPS and analysts' forecasts scaled by stock price, (winsorized at −1)." However, unlike Dhaliwal et al. (2012) he computes a simple average of the average forecast for each month 4-12 provided in the I/B/E/S summary statistics tape, whereas we include each analyst's forecast individually which in effect amounts to applying a weighted average based on the number of analysts following the firm in each month.

[19] I/B/E/S allows analysts to revise forecasts on a monthly basis, but especially for smaller companies forecasts may be revised less frequently.

Financial Information Environment ("CFIE") project has adopted the tool for usage in the corporate disclosure setting.[20] The tool is unique due to its ability to analyse unstructured corporate narrative disclosures such as those allowed under the IFRS accounting standards. Many of the studies quoted in section 2 of this paper employ different tools for electronic linguistic analysis, most notably DICTION and QSR NUD*IST, but these tools generally only work for structured corporate disclosures such as American 10-K reports. Since CSR reports like UK corporate reports generally do not follow a pre-defined format, we decided to utilise Wmatrix as our tool of analysis.

We evaluate CSR report narratives on four different aspects, which in previous research were shown to be relevant in the context of corporate reporting:

1. Readability (*READABILITY)*: Lehavy, Li & Merkley (2011) show that less readable 10-Ks are associated with greater dispersion and lower accuracy in the forecasts of financial analysts. We interpret this as indicating that more readable reports provide market participants with better information as the impact of what is reported on future performance is more easily understood by all readers (less dispersion) and also more informative (higher accuracy). Wmatrix provides both Fog and Flesch readability scores, which show a correlation of -0.96 at the 1% significance level for our sample and thus result in similar estimates for the variables in this study. Accordingly, we will show only results based on the Flesch readability scores. Our measure of CSR reporting quality is positively influenced when a report receives a higher Flesch score.
2. Length (*LENGTH)*: The length of environmental disclosures is significantly correlated with higher quality reporting (Hooks & van Staden 2011). Although a more extensive report could be used to manipulate the reader's impression by obfuscating limited actual activity, in general a longer report is likely to be an indication of more comprehensive and thus informative disclosure (Kothari, Li & Short 2009). We use word count as a measure of CSR report length, with a higher number of words having a positive effect on the narrative disclosure score.
3. Forward-looking information (*FWDINFO)*: Beattie, McInnes & Fearnley (2004) report that forward-looking information is becoming a more important attribute of disclosure quality at the expense of historical information. More extensive forward-looking information is indicative of a higher quality reporting, because it allows readers to make a better educated guess of future developments (Hussainey, Schleicher & Walker 2003).

[20] The project involved a collaboration of management researchers and NLP experts connected to the London School of Economics and Political Science, Lancaster University and the University of Manchester.

To control for the fact that longer reports are inherently more likely to contain additional occurrences of a particular word, we use the ratio of forward-looking over total words to determine how 'forward-looking' the overall report is.[21] A more forward-looking report is considered to be of higher quality.

4. Optimism (*OPTIMISM)*: Investors are susceptible to the manner in which information is formulated, interpreting the underlying message differently depending on the tone in which information is communicated (Henry 2008). Since there is no standardised framework for CSR reporting, managers are left with substantial leeway regarding the way CSR information is presented, possibly allowing them to employ some form of impression management when creating their CSR reports. Impression management frequently manifests itself through the application of an overly optimistic tone, where a high degree of optimism is likely to conceal actual performance and in turn diminish reporting quality (Cho, Roberts & Patten 2010). On the other hand, negative statements are likely to be indicators of honesty or truthfulness. Following Cho, Roberts & Patten (2010), we therefore measure the total Optimism of a report by applying lists of both positive and negative words. The list, which were developed by the Wmatrix research team, take into account the issue raised by Loughran & McDonald (2011) that many 'ordinary' negative words need not to have a negative connotation in a business context.[22] The estimated optimism of a report is obtained by subtracting the number of negative words from the number of positive words, scaled by the total amount of words. Since we believe that more honest reports contain less positive and more negative words, we argue that a lower Optimism ratio indicates better reporting quality.

In conclusion, we assume that reports are of better quality when they are easier to read, longer, contain more information about the future and do not attempt to influence the readers perception by providing overly optimistic statements.

Table 1 provides an overview of various descriptive statistics of the four quality aspects. On average, the CSR reports in our sample have a Flesch readability score of 41, which is the second highest level and indicates that readers need to have attended college to fully understand the text presented. The length of reports varies considerably, from 2 235 words to over 100 000 words. Reports can become especially long if they include an extensive GRI table next to the regular report in an attempt to comply with GRI reporting requirements. An example is the CSR report issued by Axel Springer in 2013. Reports contain between 0.7% and 7% of forward-

[21] See Appendix 1 for the forward-looking word list.
[22] See Appendices 2-5 for the negative and positive word lists.

looking words as a fraction of total words. There are only five CSR reports (untabulated) that contain more negative than positive words (resulting in a negative optimism variable), indicating that the tone is generally optimistic.

Descriptive statistics CSR report narrative characteristics

	Count	Mean	*Median*	Std. Dev.	Min	Max
READABILITY	301	40.98	*41.00*	10.11	0	105.20
LENGTH	301	25 814	*21 769*	17 668	2 235	118 868
FWDINFO	301	0.0177	*0.0161*	0.0078	0.0071	0.0714
OPTIMISM	301	0.0141	*0.0138*	0.0071	-0.0027	0.0350

Table 1: Descriptive statistics CSR report narrative characteristics

Construction of a 'Narrative Index' and grouping into three brackets

After calculating all discrete continuous variables per report for the four aspects outlined above, we rank *READABILITY*, *LENGTH* and *FWDINFO* and inverse rank *OPTIMISM* within the total sample of working CSR reports. We then sum the ranks of each variable per report to obtain a measure of a given CSR report's total narrative index, hereafter named *NINDEX*. We employ this type of ranking method for several reasons: first, it is an accepted method in disclosure studies (Lang & Lundholm 1996, Cooke 1998) and it also allows for meaningful aggregation of the four aspects. Additionally, it allows us to cope with extreme outliers within the observed data. Finally, we have no prior literature to guide us on what are 'good' versus 'bad' discrete continuous scores for each of the aspects, nor whether the relationship between the discrete continuous components of high-quality CSR reports and analyst forecast accuracy is of linear nature. Rather, this method allows us to estimate how analyst forecast accuracy changes when a report contains relatively better narrative disclosures compared to other CSR reports in the sample. Guided by Muslu et al. (2014) we subsequently split the sample into three categories representing the 'top' and 'bottom' quartile (75 observations each), and a 'middle' group consisting of the remaining 50% of observations (151 observations) to further overcome potential measurement noise. We do so based on an inspection of sorted *NINDEX* values, which indicate that scores are relatively similar across the average, but that scores in the lowest and highest quartile appear to be significantly different (Exhibit 4). While the graph suggests that the lowest and highest decile contain the most variation, we divide the values at the lowest and highest quartile in order to include sufficient observations in each group to obtain stable results in our regression analyses. Note that firm-year observations in which we find only annual reports (labelled “AR”) receive a 0 for *NINDEX* and the three reporting groups, and thus serve as the control group in all regressions.

Appendix 6 provides examples for relatively top, middle and low quality CSR reports, showing the individual scores and the corresponding relative rank as well as the resulting *NINDEX* score. Appendix 7 illustrates the distribution of top, middle and low quality CSR reports across years, countries and industries. Over the three years, the distribution is in line with the 25-50-25% split imposed by the variable construction. There are notable differences to this split across countries and industries, with especially Norway as well as the Telecommunication and Utilities industries having more higher-quality reports. Although this might indicate that reporting practices are better in Norway and these industries compared to the others, we acknowledge that this finding may also be caused by the relatively small sample of working CSR reports for these three categories.

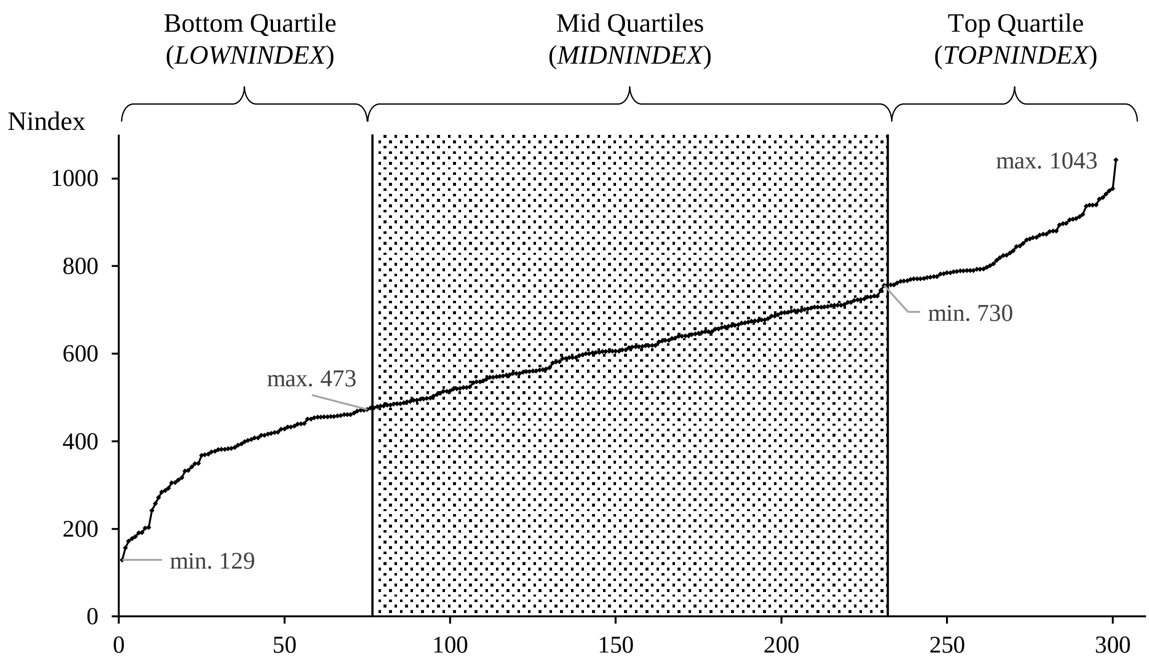


Exhibit 4: Illustration of construction of *NINDEX* groups

Validation of NINDEX

To validate that the separate components of *NINDEX* measure distinct but related aspects of disclosure quality, we analyse its constituents by looking at the correlation coefficients between the assigned ranks (Table 2). We find that the separate components are positively correlated on three aspects, but also negatively correlated on three aspects. Furthermore, four of the correlations are statistically significant. Note that the negative sign of the correlation between *FWDINFO* and *LENGTH* is probably due to the fact that the fraction (forward-looking words over total number of words) is likely to be smaller as reports become very long. The same applies to the correlation between *OPTIMISM* and *LENGTH*. However, since we argue that low *OPTIMISM* scores are a sign of better quality we apply an inverse rank and therefore observe a positive relationship. Despite this drawback it is inevitable to scale by *LENGTH* due to the inherent likelihood that a given word occurs more frequently in a longer report. It is furthermore notable that *READABILITY* is strongly and significantly correlated to *FWDINFO* and *LENGTH*, while *OPTIMISM* is negatively correlated to two other measures of disclosure quality. Despite observing negative correlations between the ranks of some of our measures of reporting quality, we nevertheless conclude that the separate constituents of *NINDEX* sufficiently measure disclosure quality since most of the negative relationships are statistically insignificant, while the signs on the positive correlations are generally larger and significant.

Correlation matrix rank *NINDEX* constituents

	READABI-LITY	LENGTH	FWDINFO	OPTIMISM
READABILITY	1.0000			
LENGTH	0.1031 *	1.0000		
FWDINFO	0.2076 ***	-0.1278 **	1.0000	
OPTIMISM	-0.0831	0.1319 **	-0.0707	1.0000

Note: *,** and *** indicate statistically significant correlations at 10%, 5% and 1%, respectively.

Table 2: Correlation matrix rank *NINDEX* constituents

To further validate the narrative index, we estimate the correlation between Asset4 scores and *NINDEX*. Asset4 scores are a measure of ESG performance based on three separate pillars (Ecological, Social and Corporate Governance) and are assigned to companies on a yearly basis by Thomson Reuters.[23] Thomson Reuters compiles Asset4 scores for around 5 000 companies globally, corresponding to 438 firm-year observations in our final sample. Appendix 8 provides an overview of the number of Asset4 (*ASSET4*) score firm-year observations per reporting type.[24] Note that only 235 of the 2 221 (10.6%) firm-year observations marked AR received Asset4 scores, while we find 203 of the 301 (67.4%) firm-year observations for which we have

[23] In essence, they appear to be similar to the more well-known KLD Sustainability scores.
[24] In this study, we do not employ the total Asset4 score but rather only take the average for the two pillars Ecological and Social, because CSR Reports do not contain information on Governance issues.

working CSR reports. We also investigate the extent to which CSR reports are externally assured across our three groups of CSR reports.

As presented in Table 3 - panel A, the average and median adjusted Asset4 score are the lowest for companies without a CSR report, and increase along the low, middle and top *NINDEX* groups. Similarly, average assurance rates increase from 27% for the lowest quality reports to 69% for the highest rated reports. We validate these results with a non-parametric test of the equality of the medians as well as a t-test for the average, and confirm that the medians and averages of the Asset4 scores for firms without a CSR report and firms within the three *NINDEX* groups are statistically different. Panel B reports similar tests for the differences between our three groups. While assurance rates are increasing and statistically different between all three groups, for Asset4 we only find statistically significant (at 10%) differences between the top and the mid group and the top and low group. Panel C reports the correlation coefficients between *NINDEX*, Asset4 and Assurance, which are all found to be strongly positively correlated. We believe that these tests provide sufficient validation of our measure of CSR reporting quality.

***NINDEX* validation**

Panel A - mean and *medians* of ASSET4 and ASSURANCE per group

	No CSR Report	LOWNINDEX	MIDNINDEX	TOPNINDEX
ASSET4	47.65	83.70 ***	84.10 ***	87.18 ***
	48.72	*88.39* ***	*88.91* ***	*91.25* ***
ASSURANCE (1)	0	0.27 ***	0.46 ***	0.69 ***
	0	*0* ***	*0* ***	*1* ***

Panel B - (statistical significance of) difference between mean and *median* of NINDEX groups

	diff. MID-LOW	diff. TOP-MID	diff. TOP-LOW
ASSET4	0.40	3.08 *	3.48 *
	0.52	*2.34* *	*2.86* *
ASSURANCE	0.19 ***	0.24 ***	0.43 ***
	0 ***	*1* ***	*1* ***

Panel C - Correlation matrix between NINDEX, Asset4 and ASSURANCE

	NINDEX	Asset4	ASSURANCE
NINDEX	1.00		
ASSET4	0.63 ***	1.00	
ASSURANCE	0.34 ***	0.41 ***	1.00

Note: *, ** and *** indicate statistically significant differences at the 10%, 5% and 1% levels, respectively, between the NINDEX groups and the no reporting group (Panel A), between the different groups (Panel B) or significant correlations (Panel C).

(1) ASSURANCE is an indicator for externally assured CSR information. Therefore, ARs by definition were assigned 0.

Table 3: *NINDEX* validation

3.4 *Empirical model*

To test our hypothesis that higher quality CSR reporting narratives is related to more accurate earnings forecasts, we adapt the baseline model developed by Dhaliwal et al. (2012) as follows:

$$FERROR = \alpha + \beta_1 TOPNINDEX + \beta_2 MIDNINDEX + \beta_3 LOWNINDEX + \beta_4 ASSURANCE + \beta_5 EARNSUP + \beta_6 LOSS + \beta_7 LEV + \beta_8 FFIN + \beta_9 ANAFOLLOW + \beta_{10} FHORIZON + \beta_{11} SIZE + \gamma\, Year * Country + \delta\, Year * Industry + \varepsilon \quad (2)$$

As discussed above, *FERROR* denotes the average absolute forecast error scaled by the opening share price. Our main variables of interest are *TOPNINDEX*, *MIDNINDEX* and *LOWNINDEX*, indicator variables corresponding to whether a firm-year observation is included in the respective group on the basis of its assigned *NINDEX* score, and thus the quality of its CSR report. Dhaliwal et al. (2012) find that the issuance of non-financial information as measured through the publication of a stand-alone CSR report reduces forecast errors. However, since they did not differentiate between CSR reports of different quality we can only with confidence predict a negative relation between *FERROR* and *TOPNINDEX*. If the findings of Dhaliwal et al. (2012) were driven by very strong reductions in forecast errors for especially good CSR reports, it could be possible to find movements in the opposite direction for average or low quality CSR reports. We therefore make no prediction as to the relation between *FERROR*, *MIDNINDEX* and *LOWNINDEX*. We also control for various additional variables, which previous studies have shown to be significant determinants of *FERROR*. Appendix 9 contains a description of each variable and its data source.

We use the gathered information on assurance statements (*ASSURANCE*) and include an indicator variable to control for whether a firm has obtained separate external assurance for its CSR report. Assurance is considered as a further legitimation of CSR work since it represents an attempt to generate more credibility among stakeholders (Simnett, Vanstraelen & Chua 2009). Hence, the signalling effect of obtaining assurance concerning the quality of the CSR information provided could imply that those forecasts improve more than those of CSR reports that are not assured. However, Dhaliwal et al. (2012) find no support for this assertion. Furthermore, since we have shown that our measure of CSR reporting quality is significantly related to assurance, we expect that our measure also captures the quality signalling effect of assurance. Therefore, we make no prediction as to the relation between *FERROR* and *ASSURANCE*.

It is more difficult to predict volatile earnings (Dichev & Tang 2009). Therefore, following Hope (2003a) we include a variable measuring the percentage change in earnings for the reported EPS compared to the prior year, *EARNSUP*. We expect *EARNSUP* to have a positive relation with *FERROR*. Furthermore, Easterwood & Nutt (1999) show that analysts tend to underreact to negative information, implying that it is likely that analysts incur greater error when forecasting the EPS of negative earnings firms. We include a dummy variable *LOSS* to capture this effect, and expect its coefficient to be positive.

The quality of the financial information provided by the company is likely to be another major determinant of the accuracy of analyst forecasts. Following Dhaliwal et al. (2012), who bases their measure on work performed by Bhattacharya, Daouk & Welker (2003), we create an indicator measure for firm-level financial transparency, *FFIN*, that assumes 1 if the company's three-year average of absolute scaled accruals is higher than the country-industry-year average, and 0 otherwise. This implies that companies with more opaque financial statements receive a 1 for *FFIN*. We expect that financial analysts find it harder to forecast the EPS of firms with more opaque financial reporting, and thus that *FFIN* is positively related to *FERROR*.

In line with Lang & Lundholm (1996), we furthermore add control variables that account for a firm's overall information environment. Firm size (*SIZE*), approximated by the natural logarithm of the opening balance of total assets, captures various factors related to the general availability of information concerning the company. Therefore, we predict a negative relation with *FERROR*. Additionally, we include leverage (*LEV*) as another indicator of information availability, arguing that firms with higher leverage are more likely to release additional information to fulfil the stringent information requirements of debtors (Ahmed & Courtis 1999). We measure leverage by dividing total liabilities by total assets. For both *SIZE* and *LEV* we use opening balances, assuming that choices regarding the amount of information supplied by a company are primarily driven by its size and leverage at the beginning of the year. Since both measures are an indication of the overall information environment, we predict a negative relation with *FERROR*, arguing that firms with more information available can be forecasted easier.

Lys & Soo (1995) suggest that as the number of analysts following a company increases, analysts will enter into a competition to enhance the accuracy of their forecasts. Hence, we include the average number of analysts following a company in a given year (*ANAFOLLOW)*, and expect a negative coefficient. Additionally, several studies find that the amount of time between the date that a forecast is published and the reporting date of the actual is significantly positively related to forecast accuracy, since more information regarding the annual profitability of a firm has become known as the year progresses (Garcia-Meca & Sanchez-Ballesta 2006). Hence, we include forecast horizon (*FHORIZON)* measuring the average (negative) number of days between the date that forecasts and the actual are issued, and expect a negative relation with *FERROR.*

Contrary to the findings presented by Hope (2003b), we neither control for cultural differences between the countries in our sample, nor for the fact that CSR reporting is mandatory in some countries (most notably Denmark) but not all. We do so for two reasons: first, we partly aim to overcome the effect of cultural differences by selecting countries that are deemed to be relatively similar. More importantly, however, is that our chosen statistical method, a fixed effects panel regression (described in more detail in section 4.5.2), is designed to eliminate the effect of time-invariant *fixed* effects that could potentially drive observed differences between companies. Since we do not expect cultural dimensions to change considerably over the three year period considered in this study, and because no country has instituted nor abolished mandatory CSR reporting during the period, we did not include controls for these variables in our regressions.

4 Empirical Results

4.1 Descriptive statistics

In Table 4 - Panel A we report descriptive statistics for the three *NINDEX* groups versus the group of firms that do not publish CSR reports. Following Dhaliwal et al. (2012), we winsorise *FERROR* at 1% to ensure that our results are not driven by extreme values.[25] We also winsorise *EARNSUP* at 5% due to the large variation that is allowed by the construction of this variable and *SIZE* at 1% since financial institutions can reach very high levels of total assets.[26]

The average forecast error is 7.23% for the non-CSR reporters and 2.85%, 1.94% and 1.50% for the low, mid and top *NINDEX* groups, respectively. We observe that CSR reporters are different from non-reporters on nearly all factors included in the analysis. Specifically, CSR reporters are less likely to publish negative earnings *(LOSS)*, seem to have less opaque financial statements (*FFIN*) and are typically followed by more financial analysts than non-reporters. The latter finding is congruent with the survey by KPMG (2015), which found that CSR

[25] We apply winsorising instead of deleting as the method of outlier treatment, since we want to avoid losing additional data points within the already small *NINDEX* groups, thereby reducing the sample of included CSR reports even further.

[26] We winsorise within the regression sample, implying that we separately winsorise the data depending on the number of observations included in the data set of the respective regression model.

reporting is most prevalent among large companies, but drops off sharply once N250 companies are considered. A table comparing all CSR reporting firms (i.e. including those with CSR reports rejected by Wmatrix) to those only publishing ARs is provided in Appendix 10. The general tenor of those statistics is similar to the one reported above.

Table 4 - Panel B shows statistical differences between the means and medians of the research variables and the three groups of CSR reporters. Although *FERROR* is not found to be statistically different across the three groups, we interpret the general decline as an indication that our hypothesis is correctly specified. Companies in either group appear to be similar across most other control variables except for *SIZE*, indicating that larger companies generally write better reports.

Appendix 11 – Panel A, B and C provide descriptive statistics of the forecast error by country, industry and year, respectively. With 3.30% the Netherlands has a notably lower average forecast error than the average of all observations, while analysts found it most difficult to forecast the profits of companies listed in Norway, where forecasts were made with on average a 9.75% error. The high forecast errors in Norway are partly explained after inspecting industry averages, as forecast errors in the sector Oil & Gas are among the largest, averaging 14.21%. We do not find this surprising given the turmoil that has hit the fossil fuel markets in recent years. Forecast errors in the Telecommunication and Utilities sectors are lower than the average at 2.21% and 3.61%, respectively. We have no clear explanation for this observation, although the small number of observations included in these sectors implies that averages could be driven by extreme values. This seems especially likely within the Utilities industry, since the median observed forecast error is much closer to the overall median. Finally, forecast errors are notably lower for the years 2013 and 2014 compared to 2012. In 2014, this is possibly the result of our sample selection, as companies that had not reported their annual results for 2015 at the time that this study was conducted are excluded.

Descriptive statistics of the final sample

Panel A - mean and *median* of variables per group						**Panel B - difference mean and *median* for reporting group**		
	no CSR report	LOWNINDEX	MIDNINDEX	TOPNINDEX	All	diff. MID-LOW	diff. TOP-MID	diff. TOP-LOW
FERROR	0.0723	0.0285 ***	0.0194 ***	0.0150 ***	0.0661	-0.0091	-0.0044	-0.0135
	0.0230	*0.0057* ***	*0.0069* ***	*0.0060* ***	*0.0196*	*0.0012*	*-0.0009*	*0.0003*
NINDEX (1)	0	372	606	833				
	0	*399*	*606*	*805*				
ASSURANCE (2)	0	0.27	0.46	0.69	0.06	0.19 ***	0.23 ***	0.42 ***
	0	*0*	*0*	*1*	*0*	*0* ***	*1* ***	*1* ***
EARNSUP	0.76	0.30 **	0.50 *	0.37 **	0.72	0.20	-0.13	0.07
	0.31	*0.17* ***	*0.21* **	*0.16* **	*0.29*	*0.04*	*-0.05*	*-0.01*
LOSS	0.29	0.07 ***	0.1 ***	0.12 ***	0.27	0.03	0.02	0.05
	0	*0* ***	*0* ***	*0* ***	*0*	*0*	*0*	*0*
LEV	0.53	0.55	0.62 ***	0.65 ***	0.54	0.07 ***	0.03	0.1 ***
	0.53	*0.58*	*0.61* ***	*0.68* ***	*0.54*	*0.03* **	*0.07*	*0.10* ***
FFIN	0.44	0.29 ***	0.33 ***	0.21 ***	0.42	0.04	-0.12 *	-0.08
	0	*0* **	*0* ***	*0* ***	*0*	*0*	*0* *	*0*
SIZE	12.24	14.83 ***	15.41 ***	16.35 ***	12.63	0.61 **	0.96 ***	1.57 ***
	12.13	*15.02* ***	*15.21* ***	*16.34* ***	*12.46*	*0.19* *	*1.13* ***	*1.32* ***
ANAFOLLOW	4.34	15.95 ***	15.69 ***	18.32 ***	5.78	-0.26	2.63 *	2.37
	2.00	*14* ***	*13* ***	*18* ***	*3*	*-1*	*5* **	*4*
FHORIZON	-178.47	-183.92	-184.28 ***	-183.76	-179.14	-0.37	-0.52	0.15
	-180	*-187* ***	*-187* ***	*-183*	*-181*	*0*	*4* **	*4* **
N	2 221	75	151	75	2 522			

Note: In panel A, *,**,*** indicate whether the reported mean or *medians* are statistically different from between the different groups and the no reporting group at the 10%,5% and 1% confidence level, respectively.

In panel B, *,**,*** indicate whether the reported mean or *medians* are statistically different from between the groups at the 10%, 5% and 1% level confidence, respectively.

(1) Tests and totals are not reported since nindex was defined to 0 for non-reporters, and devided over the three groups on the basis of rank.

(2) Tests and totals are non-informative since annual reports have no CSR report by by definition.

Table 4: Descriptive statistics of the final sample

4.2 Univariate results

Univariate correlations are reported in Table 5 (Pearson) and Table 6 (Spearman).[27] Consistent with our hypothesis, *NINDEX* is statistically negatively correlated with *FERROR*, indicating that forecast errors for companies with higher quality CSR reporting, as defined in this study, appear to be lower. Additionally, all three groups show negative and significant correlations with *FERROR*. The Pearson correlation is found to be larger for the *TOPNINDEX* and *MIDNINDEX* groups than for firms within *LOWNINDEX*, providing further indication that our hypothesis is reasonable. Of the control variables, *LOSS* has the largest correlation with *FERROR*, which is consistent with the notion that analysts are generally too optimistic in their forecasts of negative earnings firms (Hwang, Jan & Basu 1996). Additionally, *ANAFOLLOW* is strongly negatively correlated to *FERROR*, providing support for the suggestion that analysts engage in an 'accuracy competition'.

However, we note that many of the control variables are significantly correlated with each other and with the variables of interest, which might raise concerns regarding multicollinearity. Specifically, *ANAFOLLOW* and *SIZE* are strongly correlated with a co-movement coefficient of 0.70 (Pearson). This is congruent with the findings of Bhushan (1989) and seems intuitive: larger companies attract more analyst interest.[28] Given the negative correlation between both variables and *FERROR* it appears that the variables measure similar constructs concerning the information environment surrounding the company. None of the relevant absolute correlation coefficients are estimated to be above 0.8, which is sometimes suggested as an upper limit. Furthermore, only a very limited number of the correlation coefficients between control variables are found to be above (absolute) 0.5. We hence believe that our regression does not suffer from severe multicollinearity problems.

[27] The difference between the estimated Pearson and Spearman correlation coefficients lies primarily in whether the relationship between the variables is assumed to be linear (Pearson) or non-linear (Spearman). Furthermore, the Spearman correlation is usually employed when evaluating ordinal (ranked) data, which is especially relevant for our *NINDEX* variable.

[28] Bhushan (1989) measure *SIZE* as market value instead of total assets. Defining *SIZE* in similar terms, we find a correlation coefficient of 0.60 with *ANAFOLLOW* (untabulated).

Pearson correlation matrix	FERROR	TOP NINDEX	MID NINDEX	LOW NINDEX	NINDEX	CSR REPORT	ASSUR ANCE	EARN SUP	LOSS	LEV	FFIN	ANA FOLLOW	FHOR IZON	SIZE
FERROR	1.00													
TOPNINDEX (1)	**-0.06**	1.00												
MIDNINDEX (1)	**-0.08**	**-0.04**	1.00											
LOWNINDEX (1)	**-0.04**	-0.03	**-0.04**	1.00										
NINDEX (1)	**-0.11**	**0.65**	**0.66**	**0.26**	1.00									
CSRREPORT (1)	**-0.11**	**0.48**	**0.69**	**0.48**	**0.95**	1.00								
ASSURANCE	**-0.08**	**0.49**	**0.44**	**0.16**	**0.71**	**0.66**	1.00							
EARNSUP	**0.12**	-0.04	-0.03	**-0.05**	**-0.06**	**-0.07**	**-0.05**	1.00						
LOSS	**0.32**	**-0.06**	**-0.10**	**-0.08**	**-0.13**	**-0.14**	**-0.10**	**0.09**	1.00					
LEV	**0.07**	**0.07**	**0.08**	0.01	**0.11**	**0.10**	**0.09**	-0.02	**-0.05**	1.00				
FFIN	**0.11**	**-0.07**	**-0.05**	**-0.05**	**-0.10**	**-0.10**	**-0.08**	0.03	**0.15**	-0.02	1.00			
ANAFOLLOW	**-0.17**	**0.31**	**0.35**	**0.25**	**0.54**	**0.55**	**0.50**	**-0.06**	**-0.17**	**0.13**	**-0.16**	1.00		
FHORIZON	**-0.09**	-0.02	-0.03	-0.02	**-0.04**	**-0.04**	-0.03	-0.01	0.02	0.04	0.03	-0.03	1.00	
SIZE	**-0.15**	**0.29**	**0.31**	**0.17**	**0.47**	**0.47**	**0.42**	-0.02	**-0.30**	**0.31**	**-0.30**	**0.70**	**-0.05**	1.00

Note: In the correlation matrix, coefficients reported in bold indicate significance at the 5%-level.

(1) Correlation coefficients between *TOP*-, *MID*-, *LOWNINDEX*, *NINDEX* and *CSRREPORT* are irrelevant because their construction is dependent on each other.

Table 5: Pearson correlation matrix

Spearman correlation matrix

	FERROR	TOP NINDEX	MID NINDEX	LOW NINDEX	NINDEX	CSR REPORT	ASSUR ANCE	EARN SUP	LOSS	LEV	FFIN	ANA FOLLOW	FHOR IZON	SIZE
FERROR	1.00													
TOPNINDEX (1)	**-0.10**	1.00												
MIDNINDEX (1)	**-0.17**	**-0.04**	1.00											
LOWNINDEX (1)	**-0.14**	-0.03	**-0.04**	1.00										
NINDEX (1)	**-0.25**	**0.52**	**-0.68**	**0.43**	1.00									
CSRREPORT (1)	**-0.25**	**0.48**	**0.69**	**0.48**	**1.00**	1.00								
ASSURANCE	**-0.20**	**0.49**	**0.44**	**0.16**	**0.68**	**0.66**	1.00							
EARNSUP	**0.17**	-0.03	**-0.04**	**-0.05**	**-0.07**	**-0.08**	**-0.05**	1.00						
LOSS	**0.48**	**-0.06**	**-0.09**	**-0.08**	**-0.14**	**-0.14**	**-0.10**	-0.01	1.00					
LEV	**0.01**	**0.09**	**0.10**	0.02	**0.13**	**0.13**	**0.13**	-0.03	**-0.09**	1.00				
FFIN	**0.17**	**-0.07**	**-0.05**	**-0.05**	**-0.10**	**-0.10**	**-0.08**	0.04	**0.15**	**-0.06**	1.00			
ANAFOLLOW	**-0.41**	**0.22**	**0.29**	**0.20**	**0.43**	**0.43**	**0.33**	**-0.04**	**-0.20**	**0.18**	**-0.19**	1.00		
FHORIZON	0.01	-0.02	**-0.09**	**-0.05**	**-0.10**	**-0.10**	**-0.05**	-0.03	**0.05**	-0.03	0.03	**-0.08**	1.00	
SIZE	**-0.31**	**0.23**	**0.28**	**0.18**	**0.42**	**0.42**	**0.34**	-0.01	**-0.30**	**0.43**	**-0.30**	**0.72**	**-0.11**	1.00

Note — In the correlation matrix, coefficients reported in bold indicate significance at the 5%-level.

(1) Correlation coefficients between *TOP*-, *MID*-, *LOWNINDEX*, *NINDEX* and *CSRREPORT* are irrelevant because their construction is dependent on each other.

Table 6: Spearman correlation matrix

4.3 Regression results: indicative evidence that quality matters

Table 7, Model I reports the regression results to test our hypothesis, which stipulates our expectation to observe that improved CSR reporting quality has a larger effect on analyst forecast accuracy (*FERROR*). We employ a panel regression model including country-year and industry-year fixed effects to adjust for the observed differences in forecast error among different industries and countries over time.[29] Following the suggestions made by Gow, Ormazabal & Taylor (2010) we use cluster-robust standard errors at the firm level in order to mitigate problems concerning heteroscedasticity. Country-year and industry-year estimates are not reported since the coefficients are meaningless under the fixed effect model regression assumptions. We do not exclude insignificant control variables to allow for comparison to similar studies.

Compared to our treatment group of companies that do not report CSR information, we find that all three CSR reporting groups show negative estimated coefficients with forecast error. Furthermore, the estimates decrease in size from the top to the low group, consistent with our hypothesis. Companies included in the highest *NINDEX* bracket on average have a 4.53% lower forecast error than non-CSR reporters. However, this estimate is not significant at the 10% confidence level. The mid group shows a negative coefficient corresponding to a 4.09% reduction compared to non-CSR reporters, which is significant at the 10% level. Finally, companies that produce low quality CSR reports have an average forecast error that is 2.95% lower than those that publish only ARs, which is almost significant at the 10% confidence level.

Model II reports a very similar regression, but instead includes the continuous *NINDEX* score as opposed to the three reporting groups. Again, the sign is negative, in line with expectations, and the coefficient indicates that increasing *NINDEX* by 1 would result in a 0.007% increase in forecast accuracy. To put that into perspective, note that total scores fluctuate between 0 and 1043. However, this relation is also only very weakly significant with a t-statistic of -1.50 (p-value of 0.133). Estimates for the control variables are found to be very similar to those reporting in Model I. Taken together, we interpret the results of Model I and II as providing indicative evidence that our hypothesis holds some truth. Furthermore, we believe that the significance observed for *MIDNINDEX* compared to the top and low brackets is at least partly the result of the higher number of observations (151) that are included in *MIDNINDEX*.

[29] We acknowledge that controlling for firm-fixed effects would have been most appropriate (Amir et al. 2015). However, this would imply inserting 1 065 indicators representing each firm in our sample, which unfortunately is not allowed by the statistical software available to the researchers. Nevertheless, controlling for industry and country fixed effects is an established method in empirical accounting research (Gormley & Matsa 2014).

The r-square is relatively low at 7.59% indicating that there are many additional variables that explain forecast error, but it is in line with previous research (Hope 2003a, Dhaliwal et al. 2012). Nevertheless, we have identified several factors that significantly contribute to explaining the observed forecast errors. Most notably, it appears to be the case that financial analysts find it difficult to forecast the profits of loss making firms, which on average increases the forecast error by 4.08%. Additionally, the number of analysts is strongly negatively related to forecast error, as the error reduces by 0.53% for every additional analyst. Although this appears small, note that large companies typically have more than 20 analysts following them. Interestingly enough, the results show that *SIZE* is not significantly related to *FERROR*, which contradicts earlier studies (Lang & Lundholm 1996, Hope 2003a). We believe that this result is due to the very strong correlation between *SIZE* and *ANAFOLLOW*, indicating that they might measure the same underlying construct. We confirm this assertion by re-estimating the regression without *SIZE* (untabulated), and find the estimates to deviate only marginally from the ones reported above. Similar to Dhaliwal et al. (2012), *ASSURANCE* is insignificant, while our findings concerning *FFIN* contradict those reported by the authors. *LEV* has a positive sign while we expected a negative one, implying that more leverage might be related to higher *FERROR*. Despite being insignificant, we believe that the positive sign indicates that the earnings of riskier firms, as measured by leverage, are more difficult to forecast.

We also estimate the model employed by Dhaliwal et al. (2012), which only includes an indicator variable for whether or not a CSR report exists. The respective regression estimates are reported in Model III of Table 7. We increase the sample to also include firm-year observations with non-working CSR reports, in order to replicate Dhaliwal et al. (2012) as accurately as possible.[30] We continue to exclude IRs from the regression since we believe that the consolidated nature of Integrated Reports implies that they fall neither in the AR nor the stand-alone CSR category. Although Dhaliwal et al. (2012) did not do so, we believe this is due to the fact that Integrated Reporting was not officially established as a reporting style in 2007, which is the last year included in their sample. In line with the findings in that study, we identify a significant (at 10% confidence level) negative coefficient indicating that on average, firms that publish CSR reports have a 2.46% lower forecast error compared to firms that only publish annual reports. This finding reinforces the confidence that our model is specified correctly and that our results are indicative of the real underlying relationship between the quality of CSR

[30] Estimating the regression using only the primary sample (not reported) produces very similar results.

reporting and forecast errors made by financial analysts. The coefficient estimates and standard errors of the included control variables are similar to those reported in Model I and II.

Comprehensive model: Panel specification				
Dependent variable	**Pred.**	**I**	**II**	**III**
TOPNINDEX	-	-0.0453 (-1.30)		
MIDNINDEX	?	-0.0409 * (-1.71)		
LOWNINDEX	?	-0.0295 (-1.60)		
NINDEX	-		-0.00007 (-1.50)	
CSRREPORT	-			-0.0246 * (-1.82)
ASSURANCE	?	-0.0123 (-0.48)	-0.0087 (-0.38)	-0.0317 (-1.25)
EARNSUP	+	0.0073 ** (2.50)	0.0073 ** (2.49)	0.0056 ** (2.11)
LOSS	+	0.0408 *** (3.50)	0.0408 *** (3.50)	0.0436 *** (3.85)
LEV	-	0.0167 (1.26)	0.0166 (1.25)	0.0107 (0.87)
FFIN	+	-0.0034 (-0.38)	-0.0033 (-0.38)	-0.0037 (-0.47)
ANAFOLLOW	-	-0.0053 *** (-3.33)	-0.0053 *** (-3.32)	-0.0040 *** (-3.01)
FHORIZON	-	-0.0004 *** (-2.86)	-0.0004 *** (-2.86)	-0.0004 *** (-2.98)
SIZE	-	-0.0163 (-1.09)	-0.0164 (-1.09)	-0.0086 (-0.61)
Intercept		0.2161 (1.15)	0.2168 (1.16)	0.1246 (0.69)
Year and Industry fixed effects		YES	YES	YES
Year and Country fixed effects		YES	YES	YES
Number of Observations		2 522	2 522	2 818
Number of Firms		1 065	1 065	1 115
R^2		0.0759	0.0761	0.0722
Note	Dependent variable *FERROR*; values in brackets indicate t-statistics. *, ** and *** indicate statistical significance of estimated regression coefficients at the 10%, 5% and 1% confidence level, respectively.			

Table 7: Comprehensive model - Panel specification

4.4 Additional analysis: individual components of NINDEX

Having established that the data provides indicative evidence for the assertion that better quality CSR reporting narratives allow financial analysts to make more accurate profit forecasts, an interesting question remains if there are specific characteristics of quality that are especially welcomed by analysts. To investigate which of the four aspects of narrative disclosure quality that were suggested by previous research is most relevant in the context of CSR reporting, we estimate five additional regressions: one which contains a CSR report's separate index in all categories (as opposed to combining them into one *NINDEX*), and four regressions in which only the index for one category is included. We believe that it is less informative to again divide the reports into three groups due to the linear increase in the score of each separate component, and hence refrain from doing so.[31] Since all four aspects are indicators of reporting quality and we expect to observe falling forecast error along increasing quality, we predict negative coefficients for all four variables. The findings are presented in Table 8.

Model IV presents the regression which includes all four aspects of quality individually. Both *FWDINFO* and *LENGTH* are found to be negatively related to *FERROR*, even when controlling for other measures of reporting quality. Of the four *NINDEX* constituents, *FWDINFO* seems to be the most relevant aspect, reducing forecast error by 0.02% for every increase in the within-sample rank of our 301 working CSR reports. Given that financial analysts attempt to predict events that happen in the future, we are not surprised that they find forward-looking indicators most useful in refining their estimates. Furthermore, sustainability issues are generally thought of as long-term challenges, and in such a context it is plausible that forward-looking information is considered most useful. Although the estimated coefficient confirms our argument that longer reports are likely to contain more useful information for analysts, the weak statistical significance prohibits us from drawing a definite conclusion.

Over and above the other two aspects, more *READABILITY* and less *OPTIMISM* seem to contribute little to increasing the ability of financial analysts to predict future earnings. With regards to *READABILITY*, these results contradict the findings presented by Lehavy, Li & Merkley (2011), who investigated the effect of readability on forecast accuracy in the context of annual reports. We contend that this difference might arise because CSR reports are usually written for a greater audience than just financial analysts, which are the primary users of financial information. To be able to cater to this more diverse group of readers, CSR reports

[31] We realise, however, that this specification implies imposing an incremental linear relationship between increases in quality for each constituent of quality and *FERROR*, for which we have no direct theoretical support. Nevertheless, we believe that this is more informative than again splitting each variable into three indicator groups.

might in general score better on readability measures than financial reports. This implies that increased readability of CSR reports will not necessarily allow more analysts to correctly understand the information contained within the report, because most analysts are already able to interpret the information. With respect to *OPTIMISM*, the results indicate that more ‘honesty’ (as a higher *OPTIMISM* rank results from relatively more negative words compared to positive words, scaled by total words) does not appear to enable analysts to make more accurate forecasts. Insofar that *ASSURANCE* is also a measure of honesty, these insignificant findings are in line with the findings regarding *ASSURANCE* presented earlier. Synthesising the findings from model IV, it appears that especially additional forward-looking CSR information is relevant for financial analysts. On the other hand, higher readability and less optimistic disclosures seem to be less so.

Model Va to Vd depict the regressions when including only one aspect of quality at a time. We find that all aspects by themselves are negatively related to *FERROR*, although only *FWDINFO* and *LENGTH* are statistically significant. We believe that these findings suggest that even though including the four characteristics in one regression indicate that *READABILITY* and *OPTIMISM* do not provide additional explanatory power, they are nevertheless important aspects of reporting quality. Furthermore, since the quality of a report is irreversibly a combination of all aspects included in that report, we still argue that one aggregate measure is most appropriate when evaluating the effect of CSR reporting quality on analyst forecast accuracy.

Comprehensive model: *NINDEX* Constituents - Panel specification						
Dependent variable	**Pred.**	**IV**	**Va**	**Vb**	**Vc**	**Vd**
FWDINFO	-	-0.00019 *	-0.00021 *			
		(-1.79)	(-1.72)			
LENGTH	-	-0.00009		-0.00017 *		
		(-1.14)		(-1.75)		
READABILITY	-	0.000003			-0.00014	
		(0.03)			(-1.20)	
OPTIMISM	-	0.000001				-0.00011
		(0.01)				(-0.90)
ASSURANCE	?	-0.0146	-0.0228	-0.0105	-0.0203	-0.0164
		(-0.57)	(-0.78)	(-0.38)	(-0.75)	(-0.61)
EARNSUP	+	0.0073 **	0.0073 **	0.0074 **	0.0073 **	0.0074 **
		(2.50)	(2.51)	(2.51)	(2.50)	(2.51)
LOSS	+	0.0407 ***	0.0407 ***	0.0405 ***	0.0408 ***	0.0407 ***
		(3.48)	(3.49)	(3.47)	(3.50)	(3.49)
LEV	-	0.0165	0.0164	0.0167	0.0166	0.0167
		(1.24)	(1.24)	(1.26)	(1.25)	(1.26)
FFIN	+	-0.0308	-0.0031	-0.0030	-0.0033	-0.0033
		(-0.34)	(-0.35)	(-0.34)	(-0.37)	(-0.37)
ANAFOLLOW	-	-0.0053 ***	-0.0053 ***	-0.0055 ***	-0.0054 ***	-0.0054 ***
		(-3.30)	(-3.32)	(-3.44)	(-3.38)	(-3.40)
FHORIZON	-	-0.0004 ***	-0.0004 ***	-0.0004 ***	-0.0004 ***	-0.0004 ***
		(-2.86)	(-2.86)	(-2.86)	(-2.86)	(-2.86)
SIZE	-	-0.0138	-0.0164	-0.0163	-0.0164	-0.0162
		(-1.09)	(-1.09)	(-1.09)	(-1.10)	(-1.08)
Intercept		0.2177	0.2171	0.2157	0.2166	0.2134
		(1.16)	(1.16)	(1.15)	(1.16)	(1.14)
Year and Industry fixed effects		YES	YES	YES	YES	YES
Year and Country fixed effects		YES	YES	YES	YES	YES
Number of Observations		2 522	2 522	2 522	2 522	2 522
Number of Firms		1 065	1 065	1 065	1 065	1 065
R^2		0.0766	0.0764	0.0753	0.0753	0.0751
Note	Dependent variable *FERROR*; values in brackets indicate t-statistics. *, ** and *** indicate statistical significance of estimated regression coefficients at the 10%, 5% and 1% confidence level, respectively.					

Table 8: Comprehensive model: *NINDEX* constituents – Panel specification

4.5 Robustness tests

4.5.1 Testing of model assumptions

Although we are convinced that our data is best described using a fixed effects regression, since it is likely that firm level data on forecast errors is not independently and identically distributed over time or within industries and countries, we also acknowledge that our data set has limitations. Most importantly, our panel is unbalanced with the average number of observations per company being only 2.4, which is mainly due to the fact that 18% of the sample firms only have one observation (untabulated). Hence, we test various assumptions underlying our model.

First, we determine whether a random effects regression would be more suitable. This model is preferred because it is more efficient and produces unbiased slope estimators when the researcher can argue that there is no correlation between the unobserved and observed variables, but suffers from unobserved heterogeneity if this assumption proves to be false (Wooldridge 2010). We conduct a Hausman test (untabulated) and find that there is very strong evidence in favour of using a fixed effects model under the current specification of the model.

We also test the assumption that the error term has constant variance across all observations, i.e. that it is homoscedastic. We apply the test developed by Breusch & Pagan (1979) to test for linear heteroscedasticity and find that there is strong evidence for this assertion (untabulated). Assuming our model is correctly specified otherwise, this indicates that failing to adjust for heteroscedasticity could result in inefficient estimates of the standard errors. We realise that the generalised test for homoscedasticity devised by White (1980) would be more appropriate because we have no hypothesis about the form of heteroscedasticity that might be present in our data.[32] However, since the Breusch-Pagan test is in effect a special case of White's test, we believe that the outcome of this test provides sufficient evidence that the variance of the error term is not homoscedastic. To overcome this problem, we estimate all our models by clustering standard errors at firm level.

4.5.2 Robustness test: OLS model specification

An additional robustness test is performed by estimating our model under the OLS statistical specification rather than controlling for fixed effects. OLS regression requires that estimation errors are independently and identically distributed across all observations. In this setting that would imply assuming that analysts reset their profit expectations each year and actual profits are also unrelated across years. In the face of extant literature evaluating 'earnings smoothing' by especially larger companies (see for example Bartov (1993)), we believe that such an assertion is unlikely to hold true. Nevertheless, the resulting auto-correlation problem is likely to be limited within our sample, as the time dimension of our model is rather small (three years) compared to the number of unique companies (1 065). We test this assumption by performing

[32] Actually, we were unable to perform White's test because its generalised nature requires it to create k(k+1)/2 regressors, which, due to the country-year and industry-year fixed effects in our model, exceeds the maximum number of regressors allowed by the version of Stata available to the researchers.

a test for auto-correlation devised by Wooldridge (2010), and find that no strong auto-correlation is present in our data set (untabulated).

Note that although we use identical variables to facilitate comparison between the OLS and the fixed effects estimations, the procedure is in fact quite different. In an OLS specification dummies are inserted in the model to absorb country-year or industry-year effects, assuming that the underlying OLS assumptions hold true. As such, the researcher assumes that there is one 'solution' with a true intercept and consistent estimates for the coefficients and standard errors for the dummy variables that the researcher has included. When estimating a fixed effects model, on the other hand, the researcher assumes that there are unobservable individual-specific (in our case, industry-year and country-year) effects which are likely to be correlated with both the dependent and independent variables included in the model specification.[33] To erase these effects the researcher applies a 'demeaning' approach which involves subtracting the mean of all observations for a particular firm. Intercepts and the coefficient estimates for the fixed effects that the researcher wants to control for are estimated for each firm separately while optimising for a constant regression coefficient, and the utilised software package provides the average estimated constant and regression coefficients of all individual fixed effects estimators of all firms. As a result, the reported regression coefficients of the fixed effects that the researcher wants to control for are meaningless and the estimated coefficients of all variables differ from those estimated under the OLS procedure. However, the estimations no longer suffer from unobservable individual-specific effects that might drive the results found under an OLS specification. The OLS regression results are presented in Table 9.

Generally, specifying the model as an OLS has a large impact on the coefficient estimates and standard errors, implying that it is likely that there are unobservable individual specific effects that are correlated with both *FERROR* as well as the included control variables. The coefficient estimates for *TOPNINDEX*, *MIDNINDEX* and *LOWNINDEX* decrease in magnitude, with forecast errors for reports within the lowest bracket now actually receiving a positive rather than a negative sign, albeit not being statistically significant. Compared to our first model, *MIDNINDEX* is even more significantly negatively related to *FERROR*, and the size of the estimate is marginally larger than the one reported for *TOPNINDEX*. The continuous *NINDEX* is again estimated to be negatively related to *FERROR*, albeit less strongly and still insignificant. Interestingly, we find that the indicator variable for publishing a CSR report is no longer significant, thus casting some doubt on the specification of the model. In sum, however,

[33] Examples could be oil price shocks affecting primarily companies operating in the Oil & Gas industry or the outcome of general elections in a particular country.

we interpret these results as additional support for our hypothesis, namely that better quality CSR reports enable analysts to make more accurate profit forecasts.

Comprehensive model: OLS specification				
Dependent variable	**Pred.**	**I**	**II**	**III**
TOPNINDEX	-	-0.0158		
		(-1.54)		
MIDNINDEX	?	-0.0166 *		
		(-1.95)		
LOWNINDEX	?	0.0160		
		(0.99)		
NINDEX	-		-0.00002	
			(-1.24)	
CSRREPORT	-			-0.0100
				(-1.43)
ASSURANCE	?	0.0280 ***	0.0264 ***	0.0156 **
		(3.02)	(2.67)	(2.29)
EARNSUP	+	0.0087 ***	0.0087 ***	0.0070 ***
		(3.14)	(3.13)	(2.76)
LOSS	+	0.0939 ***	0.0939 ***	0.0925 ***
		(8.72)	(8.72)	(9.52)
LEV	-	0.0568 **	0.0566 **	0.0533 **
		(2.50)	(2.49)	(2.65)
FFIN	+	0.0150 **	0.0149 **	0.0124 **
		(2.20)	(2.18)	(2.11)
ANAFOLLOW	-	-0.0029 ***	-0.0028 ***	-0.0023 ***
		(-4.18)	(-4.15)	(-4.10)
FHORIZON	-	-0.0004 ***	-0.0004 ***	-0.0004 ***
		(-3.72)	(-3.72)	(-3.86)
SIZE	-	-0.0009	-0.0010	-0.0017
		(-0.28)	(-0.31)	(-0.64)
Intercept		0.0158	0.0162	0.0440
		(0.27)	(0.28)	(1.01)
Year and Industry fixed effects		YES	YES	YES
Year and Country fixed effects		YES	YES	YES
Number of Observations		2 522	2 522	2 818
Number of Firms		1 065	1 065	1 115
R^2		0.1817	0.1809	0.1925
Note	Dependent variable *FERROR*; values in brackets indicate t-statistics. *, ** and *** indicate statistical significance of estimated regression coefficients at the 10%, 5% and 1% confidence level, respectively.			

Table 9: Comprehensive model - OLS specification

Changing the specification of the model has a noticeable effect on several other variables. The impact is especially large on *ASSURANCE*, which becomes positive at 2.80% and statistically significant. This might indicate that firms seek to emphasise the credibility of their CSR reports in times of higher earnings volatility (which Dichev & Tang (2009) show to be negatively related to forecast accuracy), in turn requiring them to attract additional equity investment, which Cheng, Green & Ko (2014) found to be easier for firms with assured CSR reports. Additionally, both *LEV* and *FFIN* are now significant and appear to increase forecast errors, which was expected for *FFIN* but nor for *LEV*. As described above, the latter might indicate that analysts do not benefit from a better information environment for more levered firms but rather find it more difficult to forecast riskier firms. These additional significant variables contribute to lifting the overall explanatory power of the model, resulting in an r-square of 18.17%.

5 Limitations

5.1 Self-selection

We realise that our variable of interest, *NINDEX*, suffers from self-selection bias as it can only be constructed for firms that actually publish CSR reports, thus possibly leading us to draw incorrect conclusions regarding the relationship between forecast accuracy and CSR reporting quality. The decision to publish a voluntary CSR report is unlikely to be random, but instead a function of factors such as size, industry and measures of public scrutiny (Holder-Webb et al. 2009). Additionally, Dhaliwal et al. (2012) hypothesise that firms with generally good financial reporting might also be more likely to report on social and environmental issues. We aim to control for this self-selection problem by including control variables that capture several of these determinants, such as *FFIN*, *SIZE* and *INDUSTRY*. Yet, a more appropriate procedure to control for selection bias would be to employ the two-stage regression suggested by Heckman (1979), which includes estimating the so called inverse Mills ratio using a probit model as the first stage. However, we have opted not to undertake this endeavour due to a lack of resources and time constraints, seeing it would have involved obtaining several additional and hard to measure factors that could explain the decision to publish CSR reports.

Furthermore, selection bias is also apparent in the construction of our dependent variable, since forecast errors are only available for those firms that are followed by analysts, and whether a firm is followed or not is unlikely to be random. In fact, Bhushan (1989) shows that the decision of analysts to follow a company is driven by various factors, including ownership structure and firm size. Although it is not possible to eliminate this bias, we aim to minimise it by not excluding any company for which IBES has forecasts available. This involved manually checking whether control variables were available via different sources when there were reported as missing by either Datastream or IBES.

5.2 Controlling for the quality of financial report narratives

The quality of annual report disclosures has been shown to be linked to forecast accuracy (Lang & Lundholm 1996, Hope 2003a). Furthermore, it is possible that the quality of CSR reports and financial reports is interlinked. Hence, to avoid the risk that studying how the quality of CSR reports affects forecast accuracy of financial analysts is in fact a proxy for studying how the quality of annual reports affects those forecasts, it seems appropriate to control for the quality of annual reports. In line with previous researchers, we attempt to do so by including the control variable *FFIN* to capture the opacity of a firm's reported financials. However, we acknowledge that the *FFIN* measure does not provide information on the quality of narratives included in financial reports. One approach to filtering out the effect of financial report narratives on analyst forecast accuracy could have been to create similar firm-year *NINDEX* scores for the quality of narratives in financial reports, and subsequently include those scores as controls in our regressions. We have decided not to follow this approach because we deemed it to be outside the scope of this study due to the required additional investment in time and resources. Furthermore, we did not want to risk losing additional observations if they could not be processed by our analysis tool.

5.3 Sample selection bias

We observe further restrictions concerning our data set with regards to the exclusion of certain observations. As indicated in section 3.2.2, we collect around 15% less observations for 2014, as the respective firms did not publish their financial year 2015 earnings per share at the time that the data collection process for this study was finalised. Close inspection of the data in Exhibit 2 – Panel A reveals that especially German and Swedish firms publishing only financial reports did not report their actuals before our cut-off date. This has some effect on the nature of our dataset and contributes to an unbalanced panel. Additionally, if it were primarily smaller companies that publish their financial reports after April 1st 2016, the observed average for *FERROR* could be overstated since we found that *FERROR* is negatively related to *SIZE*.

We lose a significant number of firm-year observations in which CSR reports were published since they could not be analysed by Wmatrix. If the firm-year observations excluded were significantly different from those relating to CSR reports that remained in our sample, our study could suffer from additional sample selection bias. Therefore, we analyse the characteristics of companies that publish working and not working reports. As outlined in Table 10 we find that the average and median *FERROR*, Asset4 score, *ASSURANCE* and *SIZE* all differ significantly between firms that publish working and non-working reports.[34] We hence need to conclude that our sample of working CSR reports, which serves as the basis for the construction of *NINDEX*, is potentially biased towards better reports because they are issued by firms with better CSR performance and are more likely to be assured. Furthermore, the excluded firm-year observations appear to relate to smaller firms with higher forecast error than those that belong to included CSR reports. These differences limit the meaningfulness of *NINDEX* and the subsequent interpretation, since our sample is probably not random. However, assuming that the relation between reporting quality, Asset4 scores and assurance that we found in our working sample also holds for the total sample, we can infer that the non-working reports are likely to have been included in the lower or middle brackets. Hence, this would imply that the difference between the top, middle and lower group of reports would be even bigger than currently observed.

Difference firm-year obs for working and non-working CSR reports
Median in italics

	no NINDEX	*NINDEX*	diff.	
# of CSR Reports	296	301		
FERROR	0.0326	0.0201	-0.0125	**
	0.0095	*0.0061*	*-0.0034*	***
ASSET4	75.10	84.88	9.7819	***
	85.66	*89.23*	*3.5675*	***
SIZE	14.81	15.53	0.7129	***
	14.60	*15.39*	*0.7868*	***
ASSURANCE	0.28	0.47	0.1880	***
	0	*0*	*0*	***
Note	** or *** indicate statistical difference between the mean and median values for both groups at the 5% or 1% level, respectively.			

Table 10: Difference firm-year obs for working and non-working CSR reports

[34] Untabulated results furthermore showed that the excluded firm-year observations are not perfectly randomly distributed over the years, countries and industries.

5.4 Controlling for the information context and CSR performance

We would like to stress that the measure of CSR reporting quality that we construct is related specifically to the narratives included in those reports, i.e. written text. This implies for example that the presence of graphs in CSR reports is completely ignored. Insofar that graphs included in CSR reports contain similar useful information to those included in annual reports (DeSanctis & Jarvenpaa 1989) and those graphs are not annotated with narratives constructed in such a manner that our measures indicate those reports as being of higher quality, our measure could misclassify the usefulness of some reports. Furthermore, our measure for optimism and forward-looking information only counts occurrences of words falling in the respective categories, without evaluating the context in which those words occur. Since words are only useful within a relevant context, this could mean that some information is misclassified as improving quality.[35] Additionally, because we only focus on narratives we do not consider the 'look and feel' of the reports, which in some cases may impact decisions (Townsend & Shu 2010). Finally, we do not control for how a company scores on CSR performance (for example using Asset4 scores), because this would severely limit our sample. However, insofar that CSR performance rating agencies utilise additional information sources, or in cases where a company's reporting does not correctly (intentionally or otherwise) purport its actual performance on CSR activities, CSR performance ratings could provide additional useful information allowing analysts to make more accurate earnings forecasts.

[35] In the case of forward-looking information, irrelevant words could for example be a phrase indicating information *will* be updated if found to be incorrect.

6 Conclusion

In this study, we examine whether quality differences in CSR reporting narratives have divergent effects on the ability of financial analysts to make accurate profit forecasts. More specifically, we hypothesise that analysts make more accurate forecast for the earnings of firms that publish CSR reports containing higher quality narrative disclosures, since the disclosures contained in CSR reports currently remain largely voluntary and differ considerably. We draw upon previous literature to classify CSR disclosures of higher quality when they are more readable, longer, contain more forward-looking information and are less optimistically biased.

While controlling for various other variables that have been shown to explain forecast errors made by analysts, we find indicative evidence for our hypothesis, as the absolute size of the average estimated regression coefficients increases as we move from the group of the lowest quality reports to those included in the highest bracket. However, only the coefficient for the middle group of CSR reports is statistically significant at the 10% confidence level, while those of the low and top group have p-values above 10% but below 20%. In order to ensure the validity of our approach, we test our model under another specification and find similar indicative results. Furthermore, additional analyses reveal that in the context of CSR disclosures, especially additional forward-looking information seems to enable analysts to make more accurate earnings forecasts.

The present findings provide additional support for the efforts undertaken by various organisations to create common standards for the presentation of sustainability information, as variations in the quality of CSR disclosures are shown to impact the usefulness of CSR information to market participants. Based on the results of this study, standard setters could issue further guidance on the inclusion of more forward-looking information in CSR reports, because these disclosures appear to provide relevant and incremental information to analysts.

There has been very limited previous research into the quality of CSR reports on the scale utilised in this study. An evaluation of the extant literature revealed that this is partly due to the time-consuming nature of the most established approach for evaluating disclosure narratives, namely content analysis. We overcome this problem by relying on advances in computerised NLP techniques and previous research to efficiently and objectively score CSR report narratives using a publicly available analysis tool named Wmatrix. We verify the validity of the resulting scores by comparing them to other aspects that could potentially proxy for CSR reporting quality, and find significant correlations with Asset4 scores as well as whether external assurance was obtained. In so doing, this study additionally contributes to a developing field within accounting research that employs electronic resources to evaluate the usefulness of corporate disclosures.

Despite the promising findings presented in this study, we realise that our results remain only indicative due to the limited significance of our variables of interest. Although we argue that one of the reasons for this result could be the relatively small size of the treatment groups (301 observations divided over three groups) compared to the control group (2 221 observations), this study also suffers from several limitations. Primarily, our analysis tool was only able to evaluate just over half of the identified stand-alone CSR reports, and it appears unlikely that these exclusions are random. However, we argue that since the excluded reports tend to have lower Asset4 scores and assurance rates, it is possible that these observations would have been included in the lower or middle categories. As the observed average forecast error for the excluded companies is also higher than the forecast error of companies currently included in the middle and lower groups, we believe that the difference between especially the top and the low group could be even bigger than now observed. Additional limitations lay in the fact that we neither control for self-selection, the quality of annual report narratives or the context in which words relevant to this study are found nor for CSR performance scores. Furthermore, note that the statistical estimates reported in this study are only reliable within the present sample due to the ranking method employed. Nevertheless, the reliability of our findings within the sample is high due to the objective measurement of all studied variables. We thus believe that the tenor of our findings is generalisable to the wider universe of CSR reports due to the diligent and comprehensive nature of the study performed.

Further research could, first of all, attempt to address the limitations identified above. For example, future researchers could consider employing a different textual analysis software package such as DICTION, or work together with the researchers that developed Wmatrix to improve its accuracy for stand-alone CSR reports. In order to generalize the present results, it would be worth studying how the quality of CSR reporting narratives affects different outcome variables, such as stock price forecasts, cost of owner’s equity or access to financing. Additionally, although this study finds that especially forward-looking information appears to be most useful to financial analysts, the standard setting community might benefit from expanding our understanding of what specific types of disclosure on CSR activities are welcomed by market participants. Furthermore, since Integrated Reporting has gained a foothold in several of the countries included in this study, future researchers might consider developing a method which allows them to evaluate whether integrating sustainability and financial information has a similar effect in reducing forecast errors made by analysts.

Bibliography

Adolphs, S. 2006, *Introducing electronic text analysis: A practical guide for language and literary studies,* Routledge.

Ahmed, K. & Courtis, J.K. 1999, "Associations between corporate characteristics and disclosure levels in annual reports: a meta-analysis", *The British Accounting Review,* vol. 31, no. 1, pp. 35-61.

Alvesson, M. & Karreman, D. 2000, "Varieties of discourse: On the study of organizations through discourse analysis", *Human Relations,* vol. 53, no. 9, pp. 1125-1149.

Amir, E., Carabias, J.M., Jona, J. & Livne, G. 2016, 06/03/2016-last update, Fixed-Effects in Empirical Accounting Research, [Online]. Available from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2634089 [2016, 04/16].

Aras, G. & Crowther, D. 2009, "Corporate sustainability reporting: a study in disingenuity?", *Journal of Business Ethics,* vol. 87, no. 1, pp. 279-288.

Archel, P., Husillos, J. & Spence, C. 2011, "The institutionalisation of unaccountability: Loading the dice of Corporate Social Responsibility discourse", *Accounting, Organizations and Society,* vol. 36, no. 6, pp. 327-343.

Arena, C., Bozzolan, S. & Michelon, G. 2015, "Environmental Reporting: Transparency to Stakeholders or Stakeholder Manipulation? An Analysis of Disclosure Tone and the Role of the Board of Directors", Corporate Social Responsibility and Environmental Management, vol. 22, no. 6, pp. 346-361.

Bakar, A., Sheikh, A. & Ameer, R. 2011, "Readability of corporate social responsibility communication in Malaysia", *Corporate Social Responsibility and Environmental Management,* vol. 18, no. 1, pp. 50-60.

Barron, O.E., Kile, C.O. & O'Keefe, T.B. 1999, "MD&A Quality as Measured by the SEC and Analysts' Earnings Forecasts", *Contemporary Accounting Research,* vol. 16, no. 1, pp. 75-109.

Bartov, E. 1993, "The timing of asset sales and earnings manipulation", *Accounting Review,* vol. 68, no. 4, pp. 840-855.

Beattie, V. 2014, "Accounting narratives and the narrative turn in accounting research: Issues, theory, methodology, methods and a research framework", *The British Accounting Review,* vol. 46, no. 2, pp. 111-134.

Beattie, V. & Davison, J. 2013, "Accounting and Business Research Special Issue Call for Papers Accounting narratives: Quantification, storytelling and sensemaking", *Management Accounting Research,* vol. 24, no. 4, pp. 418.

Beattie, V., McInnes, B. & Fearnley, S. 2004, "A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes", *Accounting Forum,* vol. 28, no. 3, pp. 205-236.

Beretta, S. & Bozzolan, S. 2008, "Quality versus quantity: the case of forward-looking disclosure", *Journal of Accounting, Auditing & Finance,* vol. 23, no. 3, pp. 333-376.

Bhattacharya, U., Daouk, H. & Welker, M. 2003, "The world price of earnings opacity", *The Accounting Review,* vol. 78, no. 3, pp. 641-678.

Bhushan, R. 1989, "Firm characteristics and analyst following", *Journal of Accounting and Economics,* vol. 11, no. 2, pp. 255-274.

Botosan, C.A. 1997, "Disclosure level and the cost of equity capital", *The Accounting Review,* vol. 72, no. 3, pp. 323-349.

Brammer, S., Brooks, C. & Pavelin, S. 2006, "Corporate social performance and stock returns: UK evidence from disaggregate measures", *Financial Management,* vol. 35, no. 3, pp. 97-116.

Breusch, T.S. & Pagan, A.R. 1979, "A simple test for heteroscedasticity and random coefficient variation", *Econometrica: Journal of the Econometric Society,* vol. 47, no. 5, pp. 1287-1294.

Burke, L. & Logsdon, J.M. 1996, "How corporate social responsibility pays off", *Long Range Planning,* vol. 29, no. 4, pp. 495-502.

Cheng, B., Ioannou, I. & Serafeim, G. 2014, "Corporate social responsibility and access to finance", *Strategic Management Journal,* vol. 35, no. 1, pp. 1-23.

Cheng, M.M., Green, W.J. & Ko, J.C.W. 2014, "The impact of strategic relevance and assurance of sustainability indicators on investors' decisions", *Auditing: A Journal of Practice & Theory,* vol. 34, no. 1, pp. 131-162.

Cho, C.H., Michelon, G., Patten, D.M. & Roberts, R.W. 2015, "CSR disclosure: the more things change…?", *Accounting, Auditing & Accountability Journal,* vol. 28, no. 1, pp. 14-35.

Cho, C.H., Roberts, R.W. & Patten, D.M. 2010, "The language of US corporate environmental disclosure", *Accounting, Organizations and Society,* vol. 35, no. 4, pp. 431-443.

Clarkson, P.M., Fang, X., Li, Y. & Richardson, G. 2013, "The relevance of environmental disclosures: Are such disclosures incrementally informative?", *Journal of Accounting and Public Policy,* vol. 32, no. 5, pp. 410-431.

Clarkson, P.M., Li, Y., Richardson, G.D. & Vasvari, F.P. 2008, "Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis", *Accounting, Organizations and Society,* vol. 33, no. 4, pp. 303-327.

Cooke, T.E. 1998, "Regression analysis in accounting disclosure studies", *Accounting and Business Research,* vol. 28, no. 3, pp. 209-224.

Core, J.E. 2001, "A review of the empirical disclosure literature: discussion", *Journal of Accounting and Economics,* vol. 31, no. 1, pp. 441-456.

Cormier, D. & Magnan, M. 2014, "The impact of social responsibility disclosure and governance on financial analysts' information environment", *Corporate Governance,* vol. 14, no. 4, pp. 467-484.

Courtis, J.K. 1998, "Annual report readability variability: tests of the obfuscation hypothesis", *Accounting, Auditing & Accountability Journal,* vol. 11, no. 4, pp. 459-472.

Creswell, J.W. 2013, *Research design: Qualitative, Quantitative, and Mixed Methods Approaches,* 4th edn, Sage publications.

Davis, A.K., Piger, J.M. & Sedor, L.M. 2012, "Beyond the numbers: Measuring the information content of earnings press release language*", *Contemporary Accounting Research,* vol. 29, no. 3, pp. 845-868.

Delmas, M.A. & Burbano, V.C. 2011, "The drivers of greenwashing", *California Management Review,* vol. 54, no. 1, pp. 64-87.

DeSanctis, G. & Jarvenpaa, S.L. 1989, "Graphical presentation of accounting data for financial forecasting: An experimental investigation", *Accounting, Organizations and Society,* vol. 14, no. 5, pp. 509-525.

Dhaliwal, D.S., Li, O.Z., Tsang, A. & Yang, Y.G. 2011, "Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting", *The Accounting Review,* vol. 86, no. 1, pp. 59-100.

Dhaliwal, D.S., Radhakrishnan, S., Tsang, A. & Yang, Y.G. 2012, "Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure", *The Accounting Review,* vol. 87, no. 3, pp. 723-759.

Dichev, I.D. & Tang, V.W. 2009, "Earnings volatility and earnings predictability", *Journal of Accounting and Economics,* vol. 47, no. 1, pp. 160-181.

Dowling, J. & Pfeffer, J. 1975, "Organizational legitimacy: Social values and organizational behavior", *Pacific Sociological Review,* vol. 18, no. 1, pp. 122-136.

Easterwood, J.C. & Nutt, S.R. 1999, "Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism?", *The Journal of Finance,* vol. 54, no. 5, pp. 1777-1797.

ESG Analytics 2015, *Environmental, Social and Governance (ESG) Reporting Requirements*, [Online]. Available from: http://esganalytics.com/esg-reporting-requirements/ [2016, 02/19].

Federal Bank of St. Louis 2016, 2016-04-18-last updat*e, Number of Listed Companies per Countr*y. Available: https://research.stlouisfed.org/fred2/series/DDOM01SEA644NWDB [2016, 05/02].

Fifka, M.S. 2013, "Corporate Responsibility Reporting and its Determinants in Comparative Perspective–a Review of the Empirical Literature and a Meta-analysis", *Business Strategy and the Environment,* vol. 22, no. 1, pp. 1-35.

Fogarty, T.J. & Rogers, R.K. 2005, "Financial analysts' reports: an extended institutional theory evaluation", *Accounting, Organizations and Society,* vol. 30, no. 4, pp. 331-356.

Francis, J. & Schipper, K. 1999, "Have financial statements lost their relevance?", *Journal of Accounting Research,* vol. 37, no. 2, pp. 319-352.

Gao, F., Dong, Y., Ni, C. & Fu, R. 2015, "Determinants and Economic Consequences of Non-financial Disclosure Quality", *European Accounting Review,* [Online]. Available from: http://www.tandfonline.com/doi/full/10.1080/09638180.2015.1013049#.

Garcia-Meca, E. & Sanchez-Ballesta, J.P. 2006, "Influences on financial analyst forecast errors: A meta-analysis", *International Business Review,* vol. 15, no. 1, pp. 29-52.

Gilbert, D. & Kent, S. 2015, *BP Agrees to Pay $18.7 Billion to Settle Deepwater Horizon Oil Spill Claim*s, July 2 edn, The Wall Street Journal.

Global Institute for Partnership and Governance 2005, *Mainstreaming Responsible Investmen*t, World Economic Forum, Geneva, Switzerland.

Global Reporting Initiative 2015a, *GRI's Histor*y. Available from: https://www.globalreporting.org/information/about-gri/gri-history/Pages/GRI's%20history.aspx [2016, 02/19].

Global Reporting Initiative 2015b, *GRI Reports lis*t. Available from: https://www.globalreporting.org/services/Analysis/Reports_List/Pages/default.aspx [2016, 03/30].

Global Reporting Initiative & Accounting for Sustainability 2012, *The value of extra-financial disclosure – What investors and analysts said,* Radley Yeldar, London, United Kingdom.

Gormley, T.A. & Matsa, D.A. 2014, "Common errors: How to (and not to) control for unobserved heterogeneity", *Review of Financial Studies,* vol. 27, no. 2, pp. 617-661.

Gow, I.D., Ormazabal, G. & Taylor, D.J. 2010, "Correcting for cross-sectional and time-series dependence in accounting research", *The Accounting Review,* vol. 85, no. 2, pp. 483-512.

Gray, R., Kouhy, R. & Lavers, S. 1995, "Corporate social and environmental reporting: a review of the literature and a longitudinal study of UK disclosure", *Accounting, Auditing & Accountability Journal,* vol. 8, no. 2, pp. 47-77.

Guba, E.G. & Lincoln, Y.S. 1994, "Competing paradigms in qualitative research" in *Handbook of Qualitative Researc*h, eds. N.K. Denzin & Y.S. Lincoln, 1st edn, Sage Publications, Thousand Oaks, CA, pp. 105-117.

Guidry, R.P. & Patten, D.M. 2010, "Market reactions to the first-time issuance of corporate sustainability reports: Evidence that quality matters", *Sustainability Accounting, Management and Policy Journal,* vol. 1, no. 1, pp. 33-50.

Hackston, D. & Milne, M.J. 1996, "Some determinants of social and environmental disclosures in New Zealand companies", *Accounting, Auditing & Accountability Journal,* vol. 9, no. 1, pp. 77-108.

Harjoto, M.A. & Jo, H. 2015, "Legal vs. normative CSR: Differential impact on analyst dispersion, stock return volatility, cost of capital, and firm value", *Journal of Business Ethics,* vol. 128, no. 1, pp. 1-20.

Healy, P.M. & Palepu, K.G. 2001, "Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature", *Journal of Accounting and Economics,* vol. 31, no. 1, pp. 405-440.

Healy, P.M. & Wahlen, J.M. 1999, "A review of the earnings management literature and its implications for standard setting", *Accounting Horizons,* vol. 13, no. 4, pp. 365-383.

Heckman, J.J. 1979, "Sample selection bias as a specification error", *Econometrica,* vol. 47, no. 1, pp. 153-162.

Henry, E. 2008, "Are investors influenced by how earnings press releases are written?", *Journal of Business Communication,* vol. 45, no. 4, pp. 363-407.

Holder-Webb, L., Cohen, J.R., Nath, L. & Wood, D. 2009, "The supply of corporate social responsibility disclosures among US firms", *Journal of Business Ethics,* vol. 84, no. 4, pp. 497-527.

Holland, J.B. 1998, "Private disclosure and financial reporting", *Accounting and Business Research,* vol. 28, no. 4, pp. 255-269.

Hooks, J. & van Staden, C.J. 2011, "Evaluating environmental disclosures: The relationship between quality and extent measures", *The British Accounting Review,* vol. 43, no. 3, pp. 200-213.

Hope, O. 2003a, "Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study", *Journal of Accounting Research,* vol. 41, no. 2, pp. 235-272.

Hope, O. 2003b, "Firm-level disclosures and the relative roles of culture and legal origin", *Journal of International Financial Management & Accounting,* vol. 14, no. 3, pp. 218-248.

Hussainey, K., Schleicher, T. & Walker, M. 2003, "Undertaking large-scale disclosure studies when AIMR-FAF ratings are not available: the case of prices leading earnings", *Accounting and Business Research,* vol. 33, no. 4, pp. 275-294.

Hwang, L., Jan, C. & Basu, S. 1996, "Loss firms and analysts' earnings forecast errors", *The Journal of Financial Statement Analysis,* vol. 1, no. 2, pp. 18-30.

Hyde, K.F. 2000, "Recognising deductive processes in qualitative research", *Qualitative Market Research: An International Journal,* vol. 3, no. 2, pp. 82-90.

Initiative for Responsible Investment 2015, 2016-01-10-last update, *Global CSR Disclosure*, [Online]. Available from: http://hausercenter.org/iri/about/global-csr-disclosure-requirements [2016, 02/05].

International Integrated Reporting Council 2014, 2014-last updat*e, What? The tool for better Reportin*g, [Online]. Available from: http://integratedreporting.org/what-the-tool-for-better-reporting/ [2016, 02/19].

Ioannou, I. & Serafeim, G. 2015, "The impact of corporate social responsibility on investment recommendations: Analysts' perceptions and shifting institutional logics", *Strategic Management Journal,* vol. 36, no. 7, pp. 1053-1081.

Jones, M.J. & Shoemaker, P.A. 1994, "Accounting narratives: A review of empirical studies of content and readability", *Journal of Accounting Literature,* vol. 13, pp. 142.

Kim, Y., Park, M.S. & Wier, B. 2012, "Is earnings quality associated with corporate social responsibility?", *The Accounting Review*, vol. 87, no. 3, pp. 761-796.

Kolk, A. 2003, "Trends in sustainability reporting by the Fortune Global 250", *Business Strategy and the Environment,* vol. 12, no. 5, pp. 279-291.

Kothari, S., Li, X. & Short, J.E. 2009, "The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis", *The Accounting Review,* vol. 84, no. 5, pp. 1639-1670.

KPMG 2015, *Currents of change: The KPMG Survey of Corporate Responsibility Reporting 201*5, KPMG International, Switzerland.

Lang, M.H. & Lundholm, R.J. 1996, "Corporate disclosure policy and analyst behavior", *The Accounting Review,* vol. 71, no. 4, pp. 467-492.

Lehavy, R., Li, F. & Merkley, K. 2011, "The effect of annual report readability on analyst following and the properties of their earnings forecasts", *The Accounting Review,* vol. 86, no. 3, pp. 1087-1115.

Li, F. 2010, "The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach", *Journal of Accounting Research,* vol. 48, no. 5, pp. 1049-1102.

Li, F. 2008, "Annual report readability, current earnings, and earnings persistence", *Journal of Accounting and Economics,* vol. 45, no. 2, pp. 221-247.

Loughran, T. & McDonald, B. 2011, "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks", *The Journal of Finance,* vol. 66, no. 1, pp. 35-65.

Lys, T. & Soo, L.G. 1995, "Analysts' forecast precision as a response to competition", *Journal of Accounting, Auditing & Finance,* vol. 10, no. 4, pp. 751-765.

Merkl-Davies, D.M. & Brennan, N.M. 2007, "Discretionary disclosure strategies in corporate narratives: incremental information or impression management?", *Journal of Accounting Literature,* vol. 27, pp. 116-196.

Michelon, G., Pilonato, S. & Ricceri, F. 2015, "CSR reporting practices and the quality of disclosure: An empirical analysis", *Critical Perspectives on Accounting,* vol. 33, pp. 59-78.

Mikhail, M.B., Walther, B.R. & Willis, R.H. 1999, "Does forecast accuracy matter to security analysts?", *The Accounting Review,* vol. 74, no. 2, pp. 185-200.

Milne, M.J. & Chan, C.C. 1999, "Narrative corporate social disclosures: how much of a difference do they make to investment decision-making?", *The British Accounting Review,* vol. 31, no. 4, pp. 439-457.

Moser, D.V. & Martin, P.R. 2012, "A broader perspective on corporate social responsibility research in accounting", *The Accounting Review,* vol. 87, no. 3, pp. 797-806.

Muslu, V., Mutlu, S., Radhakrishnan, S. & Tsang, A. 2014, "Measuring Corporate Social Responsibility Report Quality Using Narratives", [Online]. Available from: http://www.bauer.uh.edu/vmuslu/ [2016, 05/12].

Neu, D., Warsame, H. & Pedwell, K. 1998, "Managing public impressions: environmental disclosures in annual reports", *Accounting, Organizations and Society,* vol. 23, no. 3, pp. 265-282.

Orlitzky, M., Schmidt, F.L. & Rynes, S.L. 2003, "Corporate social and financial performance: A meta-analysis", *Organization Studies,* vol. 24, no. 3, pp. 403-441.

Patten, D.M. 2005, "The accuracy of financial report projections of future environmental capital expenditures: a research note", *Accounting, Organizations and Society,* vol. 30, no. 5, pp. 457-468.

Patten, D.M. 2002, "The relation between environmental performance and environmental disclosure: a research note", *Accounting, Organizations and Society,* vol. 27, no. 8, pp. 763-773.

Pflugrath, G., Roebuck, P. & Simnett, R. 2011, "Impact of assurance and assurer's professional affiliation on financial analysts' assessment of credibility of corporate social responsibility information", *Auditing: A Journal of Practice & Theory,* vol. 30, no. 3, pp. 239-254.

Previts, G.J., Bricker, R.J., Robinson, T.R. & Young, S.J. 1994, "A content analysis of sell-side financial analyst company reports", *Accounting Horizons,* vol. 8, no. 2, pp. 55.

Ramnath, S., Rock, S. & Shane, P. 2008, "The financial analyst forecasting literature: A taxonomy with suggestions for further research", *International Journal of Forecasting,* vol. 24, no. 1, pp. 34-75.

Rayson, P. 2009, *Wmatrix: a web-based corpus processing environment.*, [Online]. Available: http://www.lancaster.ac.uk/people/rayson/publications/icame01.pdf [2016, 05/05].

Roca, L.C. & Searcy, C. 2012, "An analysis of indicators disclosed in corporate sustainability reports", *Journal of Cleaner Production,* vol. 20, no. 1, pp. 103-118.

Schadewitz, H. & Niskala, M. 2010, "Communication via responsibility reporting and its effect on firm value in Finland", *Corporate Social Responsibility and Environmental Management,* vol. 17, no. 2, pp. 96-106.

Schipper, K. 1991, "Analysts’ forecasts", *Accounting Horizons,* vol. 5, no. 4, pp. 105-121.

Schleicher, T. & Walker, M. 2010, "Bias in the tone of forward-looking narratives", *Accounting and Business Research,* vol. 40, no. 4, pp. 371-390.

Simnett, R., Vanstraelen, A. & Chua, W.F. 2009, "Assurance on sustainability reports: An international comparison", *The Accounting Review,* vol. 84, no. 3, pp. 937-967.

Stickel, S.E. 1992, "Reputation and performance among security analysts", *The Journal of Finance,* vol. 47, no. 5, pp. 1811-1836.

Sydserff, R. & Weetman, P. 1999, "A texture index for evaluating accounting narratives: An alternative to readability formulas", *Accounting, Auditing & Accountability Journal,* vol. 12, no. 4, pp. 459-488.

Townsend, C. & Shu, S.B. 2010, "When and how aesthetics influences financial decisions", *Journal of Consumer Psychology,* vol. 20, no. 4, pp. 452-458.

UN Global Compact 2015, *Why report*? Available from: https://www.unglobalcompact.org/participation/report [2016, 04/05].

White, H. 1980, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity", *Econometrica: Journal of the Econometric Society,* vol. 48, no. 4, pp. 817-838.

Wooldridge, J.M. 2010, *Econometric analysis of cross section and panel data,* 2nd edn, MIT press.

Appendix

Forward-looking keywords as applied by Wmatrix

accelerat#	current financial year	look forward	scope for
aim#	envisage	look ahead	scope to
anticipat#	estimat#	may	shall
await,	eventual	next	shortly
believes	expect#	novel	should
coming<0:2>year#	forecast#	optimistic	soon
coming months	forthcoming	outlook	target#
confidence	hop#	plan#	will
confident	intend	predict#	well placed
continues	intention	prospec#	well positioned
convince#	likely	remain	would
could	unlikely	renew	year# ahead

Note: # implies any following combination of letters

Appendix 1: Forward-looking keywords as applied by Wmatrix

Negative keywords as applied by Wmatrix (Table 1/3)

abando#	artifi#	closed#	critic#	delete#	disast#	down#	exagge#
abdica#	assaul#	closeo#	crucia#	delibe#	disavo#	drag#	excess#
aberra#	assert#	closin#	culpab#	delinq#	discip#	drasti#	exculp#
abetti#	attrit#	closur#	cumber#	delist#	discla#	drawba#	exoner#
abnorm#	averse#	coerce#	curtai#	demise#	disclo#	drop#	exploi#
abolis#	backda#	coerci#	cut#	demisi#	discon#	droppe#	expose#
abroga#	bad#	collap#	damage#	demoli#	discou#	droppi#	exposi#
abrupt#	bail#	collis#	damagi#	demote#	discre#	drops#	exprop#
absenc#	bailou#	collud#	dampen#	demoti#	disfav#	drough#	expuls#
absent#	balk#	collus#	danger#	denial#	disgor#	duress#	extenu#
abuse#	balked#	compla#	deadlo#	denied#	disgra#	dysfun#	fail#
abusin#	bankru#	compli#	deadwe#	denies#	dishon#	easing#	fals#
abusiv#	bans#	compul#	debarm#	denigr#	disinc#	egregi#	fatal#
accide#	barred#	concea#	debarr#	deny#	disint#	embarg#	fault#
accusa#	barrie#	conced#	deceas#	deplet#	disloy#	embarr#	fear#
accuse#	below#	concer#	deceit#	deprec#	dismal#	embezz#	fell#
accusi#	bottle#	concil#	deceiv#	depres#	dismis#	encroa#	feloni#
acquie#	boycot#	condem#	decept#	depriv#	disord#	encumb#	felony#
acquit#	breach#	condon#	declin#	dereli#	dispar#	endang#	fictit#

Note: # implies any following combination of letters

Appendix 2: Negative keywords as applied by Wmatrix (Table 1/3)

Negative keywords as applied by Wmatrix (Table 2/3)

adulte#	break#	confes#	decrea#	deroga#	displa#	enjoin#	fined#
advers#	bribe#	confin#	deepen#	destab#	dispos#	erode#	fines#
afterm#	bribin#	confis#	deeper#	destro#	dispro#	eroded#	fired#
agains#	bridge#	confli#	deepes#	destru#	disput#	erodes#	firing#
aggrav#	broken#	confro#	deface#	detain#	disqua#	erodin#	flaw#
alerte#	burden#	confus#	defama#	detent#	disreg#	erosio#	forbid#
alerti#	burned#	conspi#	defame#	deter#	disrep#	errati#	force#
aliena#	calami#	constr#	defami#	detrac#	disrup#	erred#	forcin#
allega#	cancel#	contem#	defaul#	detrim#	dissat#	erring#	forecl#
allege#	carele#	conten#	defeat#	devalu#	dissen#	errone#	forego#
allegi#	casual#	contes#	defect#	devast#	dissid#	error#	forest#
annoy#	catast#	contra#	defend#	deviat#	dissol#	errors#	forfei#
annul#	cautio#	contro#	defens#	devolv#	distor#	errs#	forger#
anomal#	cease#	convic#	defer#	diffic#	distra#	escala#	fraud#
antico#	ceasin#	correc#	defici#	dimini#	distre#	evade#	frauds#
antitr#	censur#	corrup#	defrau#	diminu#	distur#	evaded#	fraudu#
argue#	challe#	costly#	defunc#	disadv#	divers#	evades#	frivol#
argued#	charge#	counte#	degrad#	disaff#	divert#	evadin#	frustr#
arguin#	circum#	crime#	delay#	disagr#	divest#	evasio#	fugiti#
argume#	claimi#	crimin#	delaye#	disall#	divorc#	evasiv#	gratui#
arrear#	claims#	crises#	delayi#	disapp#	divulg#	evict#	grieva#
arrest#	clawba#	crisis#	delays#	disass#	doubt#	exacer#	grossl#
ground#	imprud#	insens#	lost#	neglig#	overst#	purpor#	senten#
guilty#	inabil#	insolv#	low#	nonatt#	oversu#	questi#	seriou#
halt#	inacce#	instab#	lying#	noncom#	overtl#	quit#	setbac#
hamper#	inaccu#	insubo#	malfea#	noncon#	overtu#	racket#	sever#
harass#	inacti#	insuff#	malfun#	nondis#	overva#	ration#	severe#
hardsh#	inadeq#	insurr#	malice#	nonfun#	panic#	reasse#	severi#
harm#	inadve#	intent#	malici#	nonpay#	panics#	reassi#	sharpl#
harsh#	inadvi#	interf#	malpra#	nonper#	penali#	recall#	shocke#
hazard#	inappl#	interm#	manipu#	nonpro#	penalt#	recess#	shorta#
hinder#	inappr#	interr#	markdo#	nonrec#	peril#	reckle#	shortf#
hindra#	inatte#	intimi#	misapp#	nonren#	perils#	redact#	shrink#
hostil#	incapa#	intrus#	misbra#	nuisan#	perjur#	redres#	shrunk#
hurdle#	incarc#	invali#	miscal#	nullif#	perpet#	refina#	shut#
hurt#	incide#	invest#	mischi#	object#	persis#	refusa#	slande#
hurtin#	incomp#	involu#	miscla#	obscen#	pervas#	refuse#	slippa#
idle#	inconc#	irreco#	miscon#	obsole#	petty#	refusi#	slow#
idling#	incons#	irregu#	misdat#	obstac#	picket#	reject#	sluggi#
ignore#	inconv#	irrepa#	misdem#	obstru#	plaint#	relinq#	slump#
ignori#	incorr#	irreve#	misdir#	offenc#	plea#	reluct#	smalle#

Note: # implies any following combination of letters

Appendix 3: Negative keywords as applied by Wmatrix (2/3)

Negative keywords as applied by Wmatrix (Table 3/3)

illega#	indece#	jeopar#	mishan#	offend#	pled#	renego#	solven#
illegi#	indefe#	justif#	misinf#	omissi#	poor#	renoun#	stagge#
illici#	indict#	kickba#	misint#	omit#	poses#	repara#	stagna#
illiqu#	ineffe#	knowin#	misjud#	onerou#	posing#	reposs#	stands#
imbala#	ineffi#	lack#	mislab#	opport#	postpo#	repudi#	stolen#
immatu#	inelig#	lag#	mislea#	oppose#	precip#	resign#	stoppa#
immora#	inequi#	lapse#	misled#	opposi#	preclu#	restat#	stoppe#
impair#	inevit#	lapsin#	misman#	outage#	predat#	restru#	stoppi#
impass#	inexpe#	late#	mismat#	outdat#	prejud#	retali#	stops#
impede#	inferi#	launde#	mispla#	outmod#	premat#	retrib#	strain#
impedi#	inflic#	layoff#	misrep#	overag#	pressi#	revoca#	stress#
impend#	infrac#	least#	miss#	overbu#	pretri#	revoke#	string#
impera#	infrin#	less#	mistak#	overca#	preven#	revoki#	subjec#
imperf#	inhibi#	lie#	mistri#	overch#	proble#	ridicu#	subpoe#
imperi#	inimic#	limita#	misund#	overco#	prolon#	risk#	substa#
imperm#	injunc#	linger#	misuse#	overdu#	prone#	sabota#	sue#
implic#	injure#	liquid#	misusi#	overes#	prosec#	sacrif#	suffer#
imposs#	injuri#	litiga#	monopo#	overlo#	protes#	scanda#	suing#
impoun#	injury#	lockou#	morato#	overpa#	protra#	scruti#	summon#
imprac#	inordi#	lose#	mothba#	overpr#	provok#	secrec#	surren#
impris#	inquir#	losing#	negati#	overru#	punish#	seize#	suscep#
improp#	insecu#	loss#	neglec#	oversh#	puniti#	seizin#	suspec#
suspen#	unacco#	undete#	unfulf#	unpred#	unsett#	urgenc#	warn#
suspic#	unanno#	undisc#	unfund#	unprod#	unsold#	urgent#	warned#
taint#	unanti#	undocu#	uninsu#	unprof#	unsoun#	usurio#	warnin#
tamper#	unappr#	undue#	uninte#	unqual#	unstab#	usurp#	warns#
tense#	unattr#	unduly#	unjust#	unreal#	unsubs#	usurpe#	wasted#
termin#	unauth#	unecon#	unknow#	unreas#	unsucc#	usurpi#	wastef#
testif#	unavai#	unempl#	unlawf#	unreco#	unsuit#	usurps#	wastin#
threat#	unavoi#	unethi#	unlice#	unreim#	unsure#	usury#	weak#
tighte#	unawar#	unexcu#	unliqu#	unreli#	unsusp#	vandal#	willfu#
tolera#	uncert#	unexpe#	unmark#	unreme#	unsust#	verdic#	worr#
tortuo#	uncoll#	unfair#	unmerc#	unrepo#	untena#	vetoed#	worse#
traged#	uncomp#	unfavo#	unnece#	unreso#	untime#	victim#	worsen#
tragic#	uncons#	unfeas#	unneed#	unrest#	untrut#	violat#	worst#
trauma#	uncont#	unfit#	unobta#	unsafe#	unusab#	violen#	worthl#
troubl#	uncorr#	unfore#	unoccu#	unsala#	unwant#	vitiat#	writed#
turbul#	uncove#	unfors#	unpaid#	unsale#	unwarr#	voided#	writeo#
turmoi#	undeli#	unfort#	unperf#	unsati#	unwelc#	voidin#	wrong#
unable#	under#	unfoun#	unplan#	unsavo#	unwill#	volati#	
unacce#	undesi#	unfrie#	unpopu#	unsche#	upset#	vulner#	

Note: # implies any following combination of letters

Appendix 4: Negative keywords as applied by Wmatrix (3/3)

Positive keywords as applied by Wmatrix

able#	boom#	despit#	exclus#	increa#	optimi#	record#	superi#
above#	boost#	destin#	exempl#	incred#	outper#	regain#	surpas#
abunda#	breakt#	dilige#	expand#	influe#	outsta#	resolv#	transp#
acclai#	brilli#	distin#	expans#	inform#	perfec#	revolu#	tremen#
accomp#	certai#	dream#	fantas#	ingenu#	pleasa#	reward#	unmatc#
achiev#	charit#	easier#	favora#	innova#	please#	rise#	unpara#
adequa#	collab#	easily#	favore#	insigh#	pleasu#	rising#	unsurp#
advanc#	compli#	easy#	favori#	inspir#	plenti#	rose#	upturn#
advant#	conclu#	effic#	friend#	integr#	popula#	satisf#	valuab#
allian#	conduc#	empowe#	gain#	invent#	positi#	smooth#	versat#
assure#	confid#	enable#	good#	larger#	preemi#	solid#	vibran#
assuri#	constr#	enabli#	great#	larges#	premie#	solves#	win#
attain#	courte#	encour#	grew#	leader#	presti#	solvin#	worthy#
attrac#	creati#	enhanc#	grow#	leadin#	proact#	specta#	
beat#	defini#	enjoy#	happi#	loyal#	profic#	stabil#	
beauti#	deligh#	enthus#	high#	lucrat#	profit#	stable#	
benefi#	delive#	exceed#	honor#	merito#	progre#	streng#	
best#	depend#	excel#	ideal#	more#	prospe#	strong#	
better#	desira#	except#	impres#	most#	reboun#	succee#	
bolste#	desire#	excit#	improv#	opport#	recept#	succes#	

Note: # implies any following combination of letters

Appendix 5: Positive keywords as applied by Wmatrix

Example of *NINDEX* construction for a sample of CSR reports

Company	Report Year	Country	Industry	Readability	*rank*	Wordcount	*rank*	Forward-looking info	*rank*	Optimism	*rank*	**NINDEX**
LOWNINDEX												
WOLTERS KLUWER NV	2012	Netherlands	Consumer Services	28.47	*23*	12 188	*71*	0.0100	*20*	0.0209	*43*	**157**
CRAMO OYJ	2013	Finland	Industrials	37.08	*100*	4 099	*7*	0.0122	*62*	0.0339	*3*	**172**
ASM INTERNATIONAL NV	2014	Netherlands	Technology	36.84	*96*	9 127	*36*	0.0117	*49*	0.0300	*10*	**191**
HUSQVARNA	2014	Sweden	Consumer Goods	37.43	*103*	10 139	*48*	0.0136	*95*	0.0250	*26*	**272**
BILFINGER SE	2013	Germany	Industrials	37.05	*99*	7 522	*25*	0.0121	*59*	0.0170	*101*	**284**
MIDNINDEX												
KONE OYJ	2013	Finland	Industrials	51.69	*269*	27 303	*191*	0.0084	*6*	0.0251	*24*	**490**
ADIDAS AG	2014	Germany	Consumer Goods	41.24	*158*	31 213	*214*	0.0122	*63*	0.0196	*59*	**494**
AIRBUS GROUP SE	2014	Germany	Industrials	39.23	*125*	33 725	*226*	0.0135	*88*	0.0188	*68*	**507**
G4S PLC	2012	Denmark	Industrials	35.56	*86*	18 616	*130*	0.0181	*194*	0.0171	*100*	**510**
EVONIK INDUSTRIES AG	2012	Germany	Basic Materials	34.23	*67*	46 240	*264*	0.0161	*153*	0.0166	*105*	**589**
NOKIAN TYRES PLC	2012	Finland	Consumer Goods	38.72	*122*	25 452	*178*	0.0136	*91*	0.0106	*200*	**591**
HANNOVER RUECK SE	2012	Germany	Financials	35.72	*88*	29 350	*204*	0.0090	*8*	-0.0012	*298*	**598**
ROCKWOOL INTERNATIONAL	2014	Denmark	Industrials	44.84	*209*	10 784	*56.5*	0.0157	*142*	0.0111	*194*	**600**
NOLATO AB	2013	Sweden	Industrials	40.41	*143*	14 950	*92*	0.0197	*211*	0.0059	*271*	**717**
SEB	2012	Sweden	Financials	49.09	*254*	22 274	*154*	0.0183	*198*	0.0157	*119*	**725**
TOPNINDEX												
HOCHTIEF AG	2013	Germany	Industrials	48.14	*242*	57 230	*283*	0.0219	*246*	0.0118	*182*	**953**
KONINKLIJKE BAM GROEP NV	2013	Netherlands	Industrials	68.32	*300*	49 027	*272*	0.0206	*224*	0.0134	*161*	**957**
VOLVO AB	2012	Sweden	Industrials	46.37	*228*	45 827	*263*	0.0238	*261*	0.0098	*213*	**965**
ROYAL DUTCH SHELL PLC	2012	Netherlands	Oil & Gas	52.5	*275*	31 393	*217*	0.0275	*276*	0.0100	*209*	**977**
AURIGA INDUSTRIES AS	2013	Denmark	Basic Materials	105.2	*301*	39 786	*248*	0.0184	*199*	0.0011	*295*	**1043**

Appendix 6: Example of *NINDEX* construction for a sample of CSR reports

Distribution of CSR reports across LOW-, MID- and TOPNINDEX								
	no CSR Report	CSR Reports	LOWNINDEX		MIDNINDEX		TOPNINDEX	
Year								
2012	835	96	22	23%	47	49%	27	28%
2013	781	110	26	24%	55	50%	29	26%
2014	605	95	27	28%	49	52%	19	20%
Total	2 221	301	75	25%	151	50%	75	25%
Country								
Denmark	80	35	6	17%	22	63%	7	20%
Finland	138	47	19	40%	19	40%	9	19%
Germany	896	81	22	27%	33	41%	26	32%
Netherlands	152	33	12	36%	12	36%	9	27%
Norway	385	23	1	4%	15	65%	7	30%
Sweden	570	82	15	18%	50	61%	17	21%
Total	2 221	301	75	25%	151	50%	75	25%
Industry								
Basic Materials	106	34	3	9%	21	62%	10	29%
Consumer Goods	194	48	16	33%	18	38%	14	29%
Consumer Services	190	34	5	15%	21	62%	8	24%
Financials	353	40	8	20%	22	55%	10	25%
Health Care	265	12	2	17%	10	83%	0	0%
Industrials	516	89	23	26%	50	56%	16	18%
Oil & Gas	169	10	2	20%	4	40%	4	40%
Technology	362	19	15	79%	3	16%	1	5%
Telecommunications	38	8	1	13%	2	25%	5	63%
Utilities	28	7	0	0%	0	0%	7	100%
Total	2 221	301	75	25%	151	50%	75	25%

Appendix 7: Distribution of CSR reports across LOW-, MID- and TOPNINDEX

Availability of Asset4 scores						
	Obs. with Asset4	Total observations	*% of obs*	Firms with Asset4	Total firms	*% of firms*
AR	235	2221	*10.6%*	107	902	*11.9%*
Working CSR reports	203	301	*67.4%*	107	163	*NA*
Final sample	**438**	**2 522**	***17.4%***	**214**	**1 065**	***20.1%***
Non-working CSR reports	172	296	*58.1%*	101	*91*	*NA*
IR	148	230	*64.3%*	67	108	*62.0%*
All observations	758	3 048	*24.9%*	330	1 290	*25.6%*
Note	Number of firms with working and non-working CSR reports overlap because the CSR reports of some firms do not work in all years in which they publish CSR reports.					

Appendix 8: Availability of Asset4 scores

Variable definitions

Variable	Explanation	Data Source(s)
Dependent variable		
FERROR	The per firm-year average of the absolute difference between all forecasts of and the actual earnings per share scaled by the stock price at the beginning of the year.	I/B/E/S through wrds, share price Thomson Reuters Datastream
CSR Report variables		
CSSRREPORT	An indicator variable that equals 1 if a company issued a separate CSR report during the financial year, and 0 otherwise.	GRI Report List, Company website
READABILITY	A CSR report's Flesch readability score assigned by Wmatrix. The score is calculated as Flesch = 206.835 – 1.015(total words / total sentences) – 84.6(total syllables / total words). We impose a minimum score of 0.	Wmatrix Output
LENGTH	A CSR report's total number of words.	Wmatrix Output
FWDINFO	A CSR report's total number of forward-looking words (as defined in appendix 1) divided by total number of words.	Wmatrix Output
OPTIMISM	A CSR report's total number of positive words (as defined in appendix 5) – total number of negative words (as defined in appendix 2-4) divided by total number of words.	Wmatrix Output
NINDEX	A CSR report's sum of the individual within-sample ranks for each disclosure quality category (readability, length, forward-looking information, optimism [inverse ranked]).	Wmatrix Output
TOPNINDEX	An indicator variable that assumes 1 for CSR reports with *NINDEX* scores that fall in the highest quartile (*NINDEX* greater than 730), and 0 otherwise.	Wmatrix Output
MIDNINDEX	An indicator variable that assumes 1 for CSR reports with *NINDEX* scores that fall in the middle 50% of observations (*NINDEX* between 473 and 730, inclusive), and 0 otherwise.	Wmatrix Output
LOWNINDEX	An indicator variable that assumes 1 for CSR reports with *NINDEX* scores that fall lowest quartile (*NINDEX* between 129 and 472, inclusive), and 0 otherwise.	Wmatrix Output
ASSURANCE	An indicator variable that equals 1 if assurance was obtained for a CSR report, and 0 otherwise.	CSR reports / Company websites
Control variables		
EARNSUP	The absolute difference in actual earnings between the forecasted and the previous year, over the previous year's earnings (Hope 2003a).	Thomson Reuters Datastream
LOSS	Following Hope (2003a) and Dhaliwal et al. (2012), we use an indicator variable that takes the value of 1 for loss-making firms, 0 otherwise.	Thomson Reuters Datastream

FFIN	As applied by Dhaliwal et al. (2012), we follow Bhattacharya, Daouk & Welker (2003) to derive a measure of financial transparency, defined as the three year average of scaled accruals (ACCRUAL), where ACCRUAL = (ΔCA - ΔCL - ΔCASH + ΔSTD – DEP + ΔTP) /lag(TA).[36] *FFIN* takes a value of 1 for firms with an ACCRUAL value higher than the country-industry-year mean, and 0 otherwise.	Thomson Reuters Datastream
SIZE	The natural logarithm of the total assets at the beginning of the forecasting year.	Thomson Reuters Datastream
LEV	Total liabilities over total assets at the beginning of the year.	Thomson Reuters Datastream
ANAFOLLOW	The average number of analysts following the company during the year.	I/B/E/S through wrds
FHORIZON	The average number of days (negative) between forecast publications and earnings announcement per firm year.	I/B/E/S through wrds
Other variables		
ASSET4	Average of the firm-year scores awarded in the categories Ecological and Social.	Thomson Reuters Datastream

Appendix 9: Variable definitions

[36] We calculate yearly changes for CA = Current Assets, CL = Current Liabilities, STD = Short-term portion of Debt, and TP = Taxes payable. Furthermore, we use DEP = Depreciation and Amortisation as well as TA = Total Assets.

Descriptive statistics of CSR Reports - working vs. non-working reports

Median in italics

	no CSR Report	CSR Report		Total
FERROR	0.0723	0.0205	***	0.0661
	0.0230	*0.0061*	***	*0.0196*
NINDEX	0	604		
	0	*606*		
ASSURANCE	0	0.38	***	0.08
	0	*0*	***	*0*
EARNSUP	0.76	0.41	***	0.72
	0.31	*0.19*	***	*0.29*
LOSS	0.29	0.13	***	0.26
	0.00	*0.00*	**	*0*
LEV	0.53	0.6	***	0.54
	0.53	*0.59*	*	*0.55*
FFIN	0.44	0.3	***	0.41
	0	*0*		*0*
SIZE	12.24	15.50	***	12.63
	12.13	*15.39*	***	*12.46*
ANAFOLLOW	4.34	14.81	***	6.56
	2.09	*12.17*	***	*3.00*
FHORIZON	-178.47	-183.84	***	-179.61
	-180.00	*-185.77*		*-181.42*
N	2 221	597		2 818

Note: In panel A, *,**,*** indicate whether the reported mean or *medians* are statistically different from between csr reporters and the no reporting group at the 10%,5% and 1% level, respectively.

Appendix 10: Descriptive statistics of CSR reports - working vs. non-working reports

FERROR **per year, country and industry**	Mean	*Median*	Std. Dev.	N
Year				
2012	0.0788	*0.0263*	0.1626	931
2013	0.0596	*0.0190*	0.1379	891
2014	0.0575	*0.0143*	0.1492	700
Total	0.0661	*0.0196*	0.1508	2 522
Country				
Denmark	0.0678	*0.0166*	0.1717	115
Finland	0.0471	*0.0178*	0.1295	185
Germany	0.0621	*0.0177*	0.1477	977
Netherlands	0.0330	*0.0112*	0.0742	185
Norway	0.0975	*0.0324*	0.1825	408
Sweden	0.0671	*0.0188*	0.1482	652
Total	0.0661	*0.0196*	0.1508	2 522
Industry				
Basic Materials	0.1253	*0.0420*	0.2167	140
Consumer Goods	0.0436	*0.0163*	0.0983	242
Consumer Service	0.0704	*0.0219*	0.1572	224
Financials	0.0602	*0.0198*	0.1366	393
Health Care	0.0665	*0.0197*	0.1639	277
Industrials	0.0529	*0.0170*	0.1352	605
Oil & Gas	0.1421	*0.0662*	0.2191	179
Technology	0.0553	*0.0153*	0.1314	381
Telecommunications	0.0221	*0.0075*	0.0300	46
Utilities	0.0361	*0.0195*	0.0507	35
Total	0.0661	*0.0196*	0.1508	2 522
Note	This table only shows *FERROR* for firms with working CSR report or AR.			

Appendix 11: Forecast error per country, industry and year