How to Measure Annual Report Narratives Disclosure? Empirical Evidence from Forward-Looking Information in the UK Prior the Financial Crisis

Suzan Abed

Applied Science University

Email: s_abed@asu.edu.jo

Amman-Jordan

Basil Al-Najjar*

Birkbeck University of London

Email: b.al-najjar@bbk.ac.uk

London-UK

&

Clare Roberts

Aberdeen University

Email: c.roberts@abdn.ac.uk

Forthcoming Managerial Auditing Journal

***corresponding Author**

How to Measure Annual Report Narratives Disclosure? Empirical Evidence from Forward-Looking Information in the UK Prior the Financial Crisis

Abstract

Purpose: The study aims to investigate empirically the common alternative methods of measuring annual report narratives. Five alternative methods are employed, a weighted and un-weighted disclosure index and three textual coding systems, measuring the amount of space devoted to relevant disclosures.

Design/methodology/approach: We investigate the forward looking voluntary disclosures of 30 UK non-financial companies. We employ descriptive analysis, correlation matrix, mean comparison T-test, rankings and multiple regression analysis of disclosure measures against determinants of corporate voluntary reporting.

Findings: The results reveal that while the alternative methods of forward looking voluntary disclosure are highly correlated, yet important significant differences do emerge. In particular, it appears important to measure volume rather than simply the existence or non-existence of each type of disclosure. Overall, we detect that the optimal method is content analysis by text-unit rather than by sentence.

Originality: This paper contributes to the extant literature in forward looking disclosure by reporting important differences among alternative content analyses. However, the decision regarding whether this should be a computerised or a manual content analysis appears not to be driven by differences in the resulting measures. Rather the choice is the outcome of a trade-off between the time involved in setting up coding rules for computerised analysis versus the time saved undertaking the analysis itself.

Keywords: Narrative Disclosure, Content Analysis Methods, Forward-looking Information

How to Measure Annual Report Narratives Disclosure? Empirical Evidence from Forward-Looking Information in the UK Prior the Financial Crisis

1. Overview

Disclosure is an abstract concept that cannot be measured in an unambiguous or precise manner and there is therefore a continuing debate in the literature on how to best measure disclosure and the impact of using different measures (Healy and Palepu, 2001). While a large number of alternative content analysis methods have been employed these tend to fall into one of two types, either textual content analysis or what Joseph & Taplin (2011) call disclosure abundance, which measures the amount of space devoted to relevant items and disclosure incidence or disclosure indices which instead measure the occurrence of specific items. Although there has been a debate regarding the optimal way to design disclosure indices (eg Marston & Shrives, 1991; Chow and Wong-Boren, 1987) and similar debates on how to conduct textual analysis (eg Unerman, 2000; Beattie et al., 2004; Beattie and Thomson, 2007) there has been few studies that have compared textual content analysis and disclosure indices.

Cerf (1961) is considered to be the first to use a disclosure index to evaluate the extent of disclosure, and since then, disclosure indices have been used extensively in corporate disclosure studies. Un-weighted indexes treat all disclosure items as equal, irrespective of the amount of space devoted to the topic, the number of times a topic is returned to or the importance of the topic. In contrast, a weighted index uses different weights for various disclosure items. These weights may reflect the importance of the information as seen by various types of users (Singhvi and Desai, 1971), the type of measurement, for example qualitative versus quantitative (Botosan, 1997), or the extent of disclosure (Robb et al., 2001). Other differences include the treatment of items that may not apply to all companies with some studies omitting these (e.g., Hossain et al., 2005) and others instead including them and adjusting the resultant scores for non-applicable items (e.g., Meek et al., 1995). Finally, studies vary considerably in the number of items included in the index. However, while the construction of disclosure indices involves subjective judgments, such indices are clearly valuable research instruments (Marston and Shrives, 1991; Botosan, 1997; Healy and Palepu, 2001).

Other studies have used manual content analysis to measure the extent of forward-looking information. While this may avoid the need to design a disclosure index or a list of items that are then searched for, it typically involves a choice of measurement unit. For example, words (Campbell, 2004), sentences (Milne and Adler, 1999; Deegan et al., 2000; Celik et al., 2006) and page proportions (Campbell, 2000; Gray et al., 1995) have all been used. Recently, a number of explanatory studies have employed automated content analysis to score narratives.

There is evidence that it probably does not matter if manual content analysis records disclosure in terms of pages or sentences (see for example, Hackston and Milne, 1996). However, the issue of how the results of studies seeking to explain disclosure decisions are affected by the choice of measurement metric is still far from clear. In particular, the issue of whether or not more sophisticated content analysis methods result in more useful or more valid measures of disclosure remains underexplored (Hackston and Milne, 1996).

Several studies have examined the alternative methods of disclosure in different areas. For example, Joseph and Taplin (2011) found out whether "disclosure occurrence" or "disclosure abundance" is more appropriate proxy for sustainability reporting. However, Unerman (2000) highlighted that coding by sentence or word is considered to be incomplete, since it only captures narrative disclosure which is un-useful for measuring SER. Despite this, Unerman (2000) provided a useful discussion for measuring the volume of SER disclosure in terms of proportions of pages, or paragraphs, taking into consideration non-narrative SER disclosure such as charts, tables, and pictures. On the other hand, Milne and Alder (1999) discredited coding a single word or phrase on reliability grounds; as such words and phrases have no intrinsic meanings without sentence or sentences for context. In their analysis of intellectual capital disclosures, Beattie and Thomson (2007) argued that coding by sentences appears to be straightforward if sentence coding intellectual capital refers only to one-sub-category of intellectual capital. However, different sub-categories of intellectual capital could be disclosed in the same sentence.

Previous studies of corporate disclosure practices have employed a variety of measures of FL disclosure. For example, number of studies of FL information have used disclosure indices to measure the level of FL information (Robb et al., 2001; Walker and Tsalta, 2001; Vanstraelen et al., 2003; Hossain et al., 2005 and Lim et al. 2007; Wang and Hussainey, 2013). Celik et al. (2006) and Elzahar and Hussainey (2012), instead, have manually counted the space

devoted to relevant disclosures while Hussainey et al. (2003) and Aljifri and Hussainey (2007) employ automated content analysis methods to find the extent of FL information. The main objective of the study is to seek to bring closure to some of this debate by investigating the impact of alternative quantifying methods for FL information prior to financial crisis. There have been few studies that have compared textual content analysis and disclosure indices. The current study investigates what has been ignored in previous studies, by examining the impact of method choice on the measurement of forward-looking information, and the debate about how to measure the quality of information using multi-dimensional measures (Beattie et al., 2004). Specifically, this study employs five alternative method; two using disclosure indices and three using textual analysis, measuring FL information in a sample of UK annual report narratives. These measures are not only compared to each other in terms of correlations and rankings, but each is also used in a multiple regression analysis that seeks to explain disclosure decision choices, as this is one of the most common uses of these types of disclosure measures. In the past years, both the government and the ASB have stressed the importance of narrative disclosure. Thus, the ASB has issued a series of reporting standards and statements to expand narrative disclosure. The introduction of the narrative Operating and Financial Review (OFR) is considered to be a major innovation in UK financial reporting. For example, in July 1993 the ASB issued its first statement of OFR. This statement was not an accounting standard; it was voluntary rather than mandatory and was designed as a formulation of best practice. In January 2003, the ASB issued a revised statement to reflect the later improvement in narrative disclosure. In May 2005, the ASB issued a mandatory reporting standard RS1 "Operating and Financial Review". For the first time, the disclosure of forward-looking information was considered. However, in early 2006, the government removed the statutory OFR, and the ASB issued a statement of best practice "Reporting Statement: Operating and Financial Review" which is referred to as the reporting statement or RS (OFR), that was intended to have persuasive rather than mandatory force.

This study focuses on changes in the UK regulatory environment which resulted in a change of contents and forms of the UK financial disclosure practices before the period of financial crisis.

The rest of this paper is organised in the following manner; Section 2 provides an overview of the ways in which prior studies have measured voluntary corporate disclosure; Section 3

describes the research design; Section 4 discusses, in more detail, how each of the five disclosure measures are calculated; Section 5 evaluates the alternative disclosure measures. Section 6 discusses the alternative disclosure methods. Finally, Section 7 provides a summary and an overall conclusion.

2. Literature Review

There has been some considerable debate about the use of various weighting schemes, especially in terms of their subjectivity (Ahmed and Courtis, 1999; Coy and Dixon, 2004), it has generally been thought that providing the index is 'sufficiently' extensive, it makes little difference whether a weighted or un-weighted index is employed (Chow and Wong-Boren, 1987). However, even if this is the case, one problem with the use of disclosure indices remains. They necessitate the selection of a predetermined list of items that should not only be fairly extensive, but to be capable of being used to generate replicable scores. This means they are more suitable for areas where the researcher not only knows what might be found, but can also describe this using a series of unambiguous, mutually exclusive and exhaustive descriptors, but it is less suitable for topics where disclosures might be more varied or more difficult to categorise into a number of relatively small data subjects. However, even if a relatively unambiguous and exhaustive list of possible disclosures can be developed, the problem remains that all disclosure indices, whether weighted or not, are based upon an implicit assumption that the existence of an item is more important than the amount of space devoted to it (Marston and Shrives, 1991). This may be true for some items, in particular items that can be fully reported in a simple and succinct manner, but for many items, the amount of space devoted to the disclosure is likely to be important. Firstly, it may affect the way in which the reader of the report processes the information, either because it is seen as being indicative of the importance accorded the item by the company itself, or because decisions made are subconsciously affected by the length of disclosures. Alternatively, it may simply be the case that, while information quality and disclosure length are not uniformly related to each other, lengthier disclosures will tend to contain more information. That is, using a binary system to capture the disclosure score may be misleading, as it treats a company who makes one disclosure the same as a company that makes 50 disclosures (Hackston and Milne, 1996). That is, detecting the presence or absence of an item does not capture the disclosure volume. While previous studies have used weighting to reflect the

relative importance of items, doing this also ignores the volume of disclosure. Empirical studies do not clarify the distinction between quality and quantity: it is generally assumed that the quantity of corporate disclosure has implications in determining the quality of disclosure. Previous studies have used quantity of disclosure as a proxy for quality of disclosure. Nevertheless, the questionability of this assumption is quite evident, and a debate has been commenced on the need to develop better proxy for quality (Core, 2001).

This study contributes to the findings of the previous studies by providing a comprehensive discussion of alternative disclosure indices and content analysis methods used to measure the extent of forward-looking information. As explained previously, prior studies used alternative methods to measure the extent of FL information. For example, Walker and Tsalta (2001) and Lim et al. (2007) used un-weighted disclosure indices to measure the extent of forward-looking information; they limited their disclosure indices to a few sets of items. Robb et al. (2001) and Vanstraelen et al. (2003) employed both weighted and un-weighted disclosure indices. Clarkson et al. (1999) contacted sell-side analysts to score the disclosure quality of MD&A using a survey questionnaire, the average analysts' scores were used as a measure of disclosure quality. Celik et al. (2006), Beretta and Bozzolan (2008), and Abad et al. (2008) used manual content analysis to investigate the extent of forward-looking information. Hussainey et al. (2003), Beattie et al. (2004), Aljifri and Hussainey (2007), and Athanasakou and Hussainey (2009) employed automated content analysis to investigate the extent of forward-looking information. Thus, the current study investigates what has been ignored in previous studies by examining the impact of method choice on the measurement of forward-looking information.

Because of this, various measures incorporating the volume of disclosure or textual content analysis have become more common. The manual textual content analysis would involve the researcher in reading the entire narrative report and then recording the incidence of all relevant disclosures. However, while this may avoid the need to design a disclosure index or a list of items that are then searched for, it typically involves a choice of measurement unit.

Manual textual content analysis has several advantages over computerised textual content analysis. For example, it easily permits the use of both quantitative and qualitative analyses, and it allows researchers to better interpret the meaning of specific words and phrases (Deumes, 2008). Nevertheless, it is subject to several disadvantages: it is extremely time-

consuming and therefore less cost-effective if there is a large amount of textual data (Deumes, 2008) meaning that prior studies in the field of accountancy have typically employed a small sample due to the difficulty of data collection (Beattie and Thomson, 2007). Moreover, a researcher can make mistakes and it is prone to bias, because it is very difficult to design a reliable coding techniques and classification rules (Krippendorff, 2004). Computerised content analysis is, in contrast, both more cost-effective and more flexible. However, while it has been used in previous studies to identify certain themes in texts (Abrahamson and Amir, 1996; Breton and Taffler, 2001; Beattie et al., 2004; Kothari et al., 2009: Cho et al., 2010; Davis and Tama-Sweet, 2012; Al-Najjar and Abed 2014), it has been argued that electronic word searches are not as reliable as manual reading of the annual reports (Milne and Alder, 1999; Beattie and Thomson, 2007). This may relate to the fact that to understand disclosure, the whole sentence needs to be considered, while individual word unlikely to convey much meaning (Beattie and Thomson, 2007).

A number of explanatory studies have employed completely automated content analysis to score forward looking narratives. The majority of the studies, conducted in this area, are based on the methodology of Hussainey et al. (2003). For example, Aljifri and Hussainey (2007) empirically explore the relationship between FL information and firm-specific characteristics in the emerging market of UAE while Hussainey and Al-Najjar (2011) instead explore the impact of firm and corporate governance characteristics on disclosures of FL information as well as the relationship between such disclosures and dividend policies.

3. Research Design

We employ five alternative methods of measuring disclosure are employed. Two of these use the same pre-identified list of items, with one being un-weighted and one weighted index. However, whether a disclosure index is weighted or un-weighted, it ignores the volume of disclosure and treats a company who makes one limited disclosure the same as a company that makes very extensive disclosures. Three textual content analysis measures were therefore also employed. These instead search for each instance of disclosure and then count the space devoted to all relevant disclosures. The first of these manually counted the number of relevant text units, then the computer program QSR NVivo 8, has been employed to gain code and count relevant text units as adopted by Beattie et al. (2004) and Beattie and

Thomson (2007) and finally, QSR NUD*IST 6[1] is employed to perform coding by sentences as used by Hussainey et al. (2003) and Hussainey and Walker (2009). These five methods are then used to investigate the voluntary disclosure of forward-looking information in the narrative disclosures[2] in the 2006 annual report of 30 UK companies[3]. The sample is randomly chosen from the largest UK listed firms by market capitalization as listed by Financial Times on 30 March 2007, from eight different industries[4]. In order for the firms to be included in our models their annual narrative should be at least 50 pages. Several sets of analyses are employed to compare the resulting scores. Firstly, they are compared using correlation analysis to gain information on how similar or dissimilar they are. However, the correlation is a measure of the overall level of agreement between two measures and it says nothing about how any differences in relative scores of the sample companies are distributed. Therefore, the scores for each method are also converted into ranks and these are compared by a company to company basis to explore the issue of whether differences in the ranks are randomly distributed across the sample or whether instead some, or all, of the methods generate very similar results for some companies, but very different results for the other ones. Regression analyses are then used to gain some insight into the potential impact of using the different disclosure measures in empirical studies, to provide evidence on whether or not the results of prior empirical studies might depend crucially upon the choice of disclosure metric. Hence, our main question in this study is: Are alternative methods of measurements providing quite similar inferences of FL disclosure?

4. Disclosure Measures

In order to create a comprehensive list of disclosure items, previous studies on FL information disclosure were reviewed (e.g., Cahan and Hossain, 1996; Robb et al., 2001; Walker and Tsalta, 2001; Vanstraelen et al., 2003; Hossain et al., 2005) as was Accounting

[1] QSR NVivo 8 supersedes N6; it includes the features of N6, as well as the ability to work with PDF documents, photos, videos, and audio files.

[2] For all methods, all images, charts, tables, pictures, and graphics were ignored, as were the financial statements, notes and audit report and any narrative sections which are considered to be highly standardised, i.e. directors' reports, corporate governance report, remuneration report, list of content, list of directors and advisors, statement of directors' responsibilities, financial history and financial summary, and shareholder information. This is consistent with previous studies (e.g., Walker and Tsalta, 2001; Hussainey et al., 2003; Beattie et al., 2004) and leaves the narrative, largely voluntary sections, such as, chairman's, CEO's and financial director's statements, operating and financial review, business review or review of the operation, community and people, and environmental report.

[3] 2006 was the first year that the requirement to follow the ASB's Reporting Statement 1 : Operating & financial review, was removed.

[4]Financial companies were excluded because the very different activities and market characteristics of this industry suggest that they may consistently report different levels or types of FL information.

Standard Board (ASB) RS1 Operating and Financial Review (2005), resulting in a preliminary list of 54 items. This was then pilot-tested on 10 UK non-financial companies randomly selected to ensure that all relevant items are included and irrelevant or misleading items are excluded from the final list[5]. This process has resulted in a final list of all the original 54 items plus 10 additional items that were also disclosed by the pilot companies. These 64 items were classified into two main categories: non-financial and financial items. The former category comprises four sub-groups: (a) the market and other external information (10 items), (b) company trends (9 items), (c) strategic information (8 items), and (d) other information (8 items). The financial category consists of 29 items.

The un-weighted index was generated using a dichotomous approach; if the company discloses a specific item at least once, it is scored as 1, and 0 otherwise. Following Haniffa and Cooke (2005) a relative "Disclosure Score Ratio" (DSR) was then calculated for each company by converting the absolute score into a percentage score. While some items (for example backlog/new orders and R&D) might be expected to be more often disclosed in some industries than others, the index contained no items that would necessarily be irrelevant for a particular industry or company included in the sample. Therefore, the DSR was calculated on the basis that the maximum possible score for all companies was 64. To gain further insight into the issue of whether or not the weighting system chosen is important, one weighted disclosure score was also employed. Based upon the argument that there is more information contained in a quantitative than a qualitative forward looking disclosure item, quantitative items were more highly weighted. Following Botosan (1997), Bozzolan, (2003), quantitative items were assigned a weight of 2 and qualitative items a weight of 1[6]. Again, the absolute scores were converted to percentages.

Textual content analysis requires selection of a unit of analysis, or "*the specific segment of content that is characterised by placing it into a given category"* (Holsti, 1969, P. 116). There has been debate regarding the use of words, sentences, or paragraphs as a basis for coding. The earlier content analysis studies were generally concerned with social and environmental disclosures, and most commonly used sentences as a basis for coding decisions (Gray et al., 1995). Sentences were used because, as Milne and Adler (1999)

[5] These are a different set of companies from the 30 employed in the main analysis.
[6] A common alternative approach is to instead ask users what weights they would employ (eg Chow & Wong-Boren 1987) but this is still subjective, as it not clear which user groups should be consulted.

argued, sentences were seen to be the most reliable unit of analysis, being more likely to provide complete, reliable, and meaningful data for further analysis in contrast to coding on the basis of a single word or phrase. However, Beattie et al. (2004) argue that sentences can often be too large a unit of code, and that it is necessary to split sentences into text units, or a group of words containing a single "piece of information that was meaningful in its own right" (p. 216). For example, in the analysis of forward-looking information, using the text unit as a unit of analysis results in coding the sentence: *Despite continuing high raw material prices we are optimistic that our Polymer business will once again deliver strong results* (*Yule Catto, 2006 Annual reports and Accounts*) into two text units, one about likely material costs and one about future results. Given this disagreement over the most appropriate unit of analysis, the current study, therefore, uses both sentences and text units.

Once the content has been coded on the basis of text units, quantification can be achieved in a number of ways, the most common being by word, sentence or proportion of the page. The choice is important as the accuracy of recording is likely to vary across the alternative measures, suggesting that sentences may be the most objective measure (Hackston and Milne, 1996). However, Unerman (2000) argue that using sentences as a unit of measurement ignores the possibility that differences in sentence length may lead to very different scores for companies that in practice disclose the same amount of information. While Unerman (2000) suggest measuring the proportion of pages as a solution to grammatical differences, using text units would also seem to solve the problem of how to cope with systematic differences in sentence structures. Therefore, again sentences and text units are both used in the computerised content analysis. Thus, the manual content analysis was a fully manual one, involving reading the annual report narratives and then counting the number of text units that provide FL information, while QSR NVivo 8 software was used to conduct computerised content analysis using the text unit both as the unit of analysis and unit of quantification and N6 (NUD*IST) was used for a similar analysis using sentences for both unit of analysis and unit of measurement (Al-Najjar and Abed, 2014).

In order to perform the automated content analysis of NVivo 8, a list of FL keywords and a list of topic keywords was generated and the software used to count the occurrence of at least one FL keyword associated with a topic keyword by searching for the intersection of the two. Hussainey et al.'s (2003) list was used as a preliminary list of keywords. This was then modified and potential keywords added based upon the manual content analysis (for example,

"possible", "probable", "opportunity", "further", "chance", "future", "commitment", "near term" and "medium term") while others were instead excluded due to their very low frequency of occurrence (for example, "envisage", "eventual", "novel", "scope for", "scope to", and "shortly"). The final list of 36 forward looking words is shown in Table 1.

INSERT TABLE 1 HERE.

The selection of disclosure topics was then undertaken using four steps. Firstly, a list of topic keywords was generated based on previous textual content analysis studies (Breton and Taffler, 2001; Hussainey et al., 2003; Beattie et al., 2004), disclosure index studies (Robb et al., 2001; Walker and Tsalta, 2001; Vanstraelen et al., 2003), or frequency count studies (Cahan and Hossain, 1996; Hossain et al., 2005; Bozzolan and Mazzola, 2007). Then, when carrying out the manual content analysis, forward-looking statements were identified and any additional main topic keywords found were added to the preliminary list. This resulted in a modified list of topics which were then classified into four main themes, as employed by a number of prior studies (Robb et al., 2001; Hutton et al., 2003; Vanstraelen et al., 2003; Bozzolan and Ipino, 2007).

This list includes: *financial information*: sales, profit, cost, asset, liability, equity, capital, and cash; *strategy information*: mission, policy, and objective; *structure information*: expansion, development, improvement, growth, merger, and progress; *corporate environment information*: legal, politics, technology, competition, and risk. Again, synonyms were then added and text search-queries used to assess the applicability the list. Table 2 presents the dominant themes and the main keywords related to these themes.

Again, synonyms were then added and text search-queries used to assess the applicability the list. Table 2 presents the dominant themes and the main keywords related to these themes.

INSERT TABLE 2 HERE.

Before importing the annual reports into NVivo, the annual reports were transferred from PDF files into standard text format files. All images, charts, tables, pictures, and graphics were therefore automatically deleted. The financial statements, notes and audit report were then deleted from the text file as were those narrative sections which are considered to be

highly standardised, i.e. directors' reports, corporate governance report, remuneration report, list of content, list of directors and advisors, statement of directors' responsibilities, financial history and financial summary, and shareholder information. This is consistent with previous studies (e.g., Walker and Tsalta, 2001; Hussainey et al., 2003; Beattie et al., 2004) and leaves the narrative, largely voluntary section, such as, chairman's, CEO's and financial director's statements, operating and financial review, business review or review of the operation, community and people, and environmental report.

Text search queries were used to find the recurrence of a particular word or phrase or group of keywords in the narratives. To do this, the forward-looking keywords were classified into four groups. Keywords involving plural or singular forms, such as "chance" and "chances"; where the use of the question sign (?) after "chance?" enables the software to capture both forms. The second group is phrases consisting of two to four words, for example, "coming year" or "coming financial year". In this case, the double quotes (" ") mean the phrase is searched for. The third group is verbs, such as "forecast". However, since verbs may sometimes launch information related to the past, different forms derived from the verbs are used to reduce noise that may otherwise be introduced. These forms include "forecast", "forecasts", "is forecasted", "are forecasted", "is forecasting" and "are forecasting". Hence, they exclude past tenses such as "forecasted", "was forecasting", "were forecasting", "has forecasted", "have forecasted", "was forecasted" and "were forecasted" (e.g., Hussainey et al., 2003). Furthermore, in some cases an adverb and/or (not) distinguishes between the auxiliary and the main verb; an example of this would be "is not also forecasted". In this case, NVivo provides an option to use the Tilde (~) to indicate proximity (e.g., Wasley and Wu, 2006). As a result, the following text search query is written: "is forecasted"~2, where 2 represent the number of words between the auxiliary and the main verb. The fourth group is differentiated by the occurrence of a preposition before the year number; the text search queries which are written to find a year number should be preceded by one of the following prepositions "in", "into", "for", "of", after", "before", "through", "throughout", "by", and "during" and also followed by a year number from 2007 successively until 2030. For example, the following text search query is written to search for year number "[of|for|into|in|during|through|throughout|by|after|before][2007|2008|2009|2001|2011|2012|2013|2014|2015|2015|2016|2018|2019|2020|2021|2022|2023|2024|2025|2026|2027|2028|2029|2030]". Based on the above, a number of text search queries are written to search for FL keywords and similar procedures were applied to search for topics keywords.

An advanced feature of NVivo "coding query" was then used to find the intersection between forward-looking keywords nodes (Table 1) and topics nodes (Table 2). In an advanced coding query, more criteria should be identified to find the intersection between the previous nodes. This is performed using the option to search for text that is in close proximity by employing the "NEAR *content"* feature. The option of "Specified Number of Words" was used with seven words being chosen as the basis for coding text units (e.g., Wasley and Wu, 2006).

N6 (NUD*IST) was used to generate the final measure, disclosure measured by sentence, where the vector node feature was used to count all sentences containing at least one of the same set of forward looking words and one of the same set of relevant topics.

QSR NVivo 8 was realised in March 2008[7] as a replacement for the earlier N6 (NUD*IST). However, the "command file" in N6 has the same characteristics as "query" in NVivo 8, but it performs coding at the level of sentences. The forward-looking keywords and the topics keywords are the same as used for NVivo, but sentences are both the unit of analysis and the unit of measurement. A special feature in N6 is a vector node (which is equivalent to advanced text queries in NVivo) which is used to find all sentences containing at least one forward-looking keyword and one topic. Table 3 presents the command file to search for "predict"; the left column represents the command and the right column represents the definition of the command components.

INSERT TABLE 3

5. Evaluation of the Alternative Disclosure Measures

Table 4 provides details of the scores for each disclosure method. Firstly, it appears that there is very little difference between the results using a weighted or an un-weighted index, supporting earlier findings (Chow and Wong-Boren, 1987). However, the weighted index generates a slightly smaller mean, (32 vs. 37%) as well as slightly smaller median, standard deviation, minimum and maximum, indicating that companies tend to bias their disclosures towards qualitative information.

[7] The QSR NVivo 8 home page is at http://www.qsrinternational.com.

INSERT TABLE 4

As might also be expected, automated content analysis using NVivo software results in much higher scores than those generated by NUD*IST (mean 71.83 vs. 48.13) suggesting that there are a significant number of sentences that contain more than one relevant forward looking item. Both NVivo and the manual content analysis use text units as a unit of analysis, suggesting that the two should be similar. However, they yield quite dissimilar results, with the manual method yielding a much lower mean score (54.87 vs. 71.83), and a maximum score (122 vs. 197) but a higher minimum score (25 vs. 16). As expected, the manual method of counting text units results in a higher score than recorded by NUD*IST which records potentially longer and less frequent sentences, although the difference is not so great (54.87 vs. 48.13) and again the manual method records a higher minimum score (25 vs. 15) and a lower maximum score (122 vs. 137).

However, even the distributions are quite similar; they may mask significant differences in the scores recorded for individual companies if on average such differences cancel each other out. To examine this possibility, the Pearson correlations are reported in Table 5.

INSERT TABLE 5

The correlation matrix reports significant and relatively high correlations across all five methods. The two highest correlations are between the weighted and un-weighted disclosure indices (0.989), suggesting that companies that disclose more quantitative measures also disclose more qualitative ones, and between NVivo and NUD*IST (0.922), indicating that it may matter little if sentences or text units are recorded. However, all correlations are high, with the lowest being between NVivo (number of the text-units) and the un-weighted and weighted disclosure indices (0.810 and 0.817 respectively). All of the correlations between the indices and the textual measures fall somewhere between those found for the log of the number of items and sentences by Jospeh and Taplin (2011) at 0.78 and the 0.93 (in 1996 or 0.94 in 2002) found by Haniffa and Cooke (2006) for the number of words and the number of items. Joseph and Taplin (2011) argue that the correlation will depend crucially upon the number of items reported, with the correlation increasing as the amount of disclosure falls. However, this is at best only a partial explanation of the differences, given that while the

mean number of items found by Joseph and Taplin (2011) was 15.3 and Haniffa and Cooke (2006) report an average of 6.9 items (in 2002), this study in contrast found a higher level of disclosure than either of these studies, with the mean number of items disclosed being 23.7. While all the correlations are high, they are not perfect and it may be that the differences between them are important. They may be important if, for at least some companies, the resultant conclusions drawn about the company differs according to the measures being considered. In order to gain more understanding about the differences between the five different methods used we introduce a Mean comparison T-test between the weighted and unwegited methods and the three different textual coding systems and report the results in Table 6. We detect that all the comparisons, apart from the manual method and the NUD*IST, are significant, indicating there are significant differences among these methods.

INSERT TABLE 6

To explore why these differences emerge, the disclosure scores for each method were replaced by their corresponding ranks (30 being the highest disclosing company) and for each company the standard deviation of the five resultant ranks was also calculated, as reported in Table 7. This shows that, on average, the standard deviation of the five ranks is 2.96. However, there are a few large differences across the 30 companies, the standard deviations for each company varying from 0 (GKN) to 8.10 (Premier Oil).

INSERT TABLE 7

Again, there are very few differences between the weighted and un-weighted indices. They both agree on the three best companies (GKN, Scottish and Newcastle and British Airways) while 20 of the 30 companies were within two ranks of each other. Indeed, the maximum difference in rank was only 4.5 with *Uniq* disclosing relatively more quantitative information and less qualitative information and *WSP* and *Yule Catto* producing instead relatively more qualitative disclosures but less quantitative information than the average. Looking at NUD*IST and NVivo, seven companies have the same rank, these including the best three (now *GKN, British Airways and Smiths group*) and the poorest two (Hill & Smith, JJB).

Another 12 companies have a difference in rank of no more than three. In contrast, there are three companies that yield very different results. WSP (9 NVivo and 17 NUD*IST) and Melrose Resources (4 vs. 21) appear to be unusual with simpler sentences than the average company, while in contrast, Premier Oil (26 vs. 16) appears to have a more complex sentence structure than the average company. When all five methods are compared, all agree that the highest level of disclosure is found in GKN's report and they also all tend to agree that British Airways is particularly good, while Hill & Smith produces the least information across four measures and second lowest only for unweighted index, with Greggs also being particularly poor. 15 of the 30 companies have a range of ranks of five or less, but in contrast a few very large differences appear across the five methods for some of the other companies. However, there appears not to be any discernible pattern in these differences. For example, for Melrose the largest difference is between NVivo and NUD*IST, (4 vs. 21). In contrast, *InterCon*, which has the second highest standard deviation of all companies, yields virtually identical ranks for NVivo and NUD*IST (22 and 20) but very different ranks for the frequency and manual content methods (4 and 3). In contrast, Premier Oil, with the highest standard deviation, yields ranks under the disclosure indices that are out of line with the other three methods as well as very different ranks for the two automated content analysis methods. However, while there are no clear patterns of differences two conclusions can be drawn from this table, firstly, at least for the weighting scheme used, there are few differences in the ranks of the two disclosure indices. Secondly, while some companies yield consistent ranks across all five methods, a number of companies instead yield quite large differences in their ranks across two or more of the three textual content analysis methods or between one or more of these three methods and the two disclosure indices.

Given that the different disclosure measures yield quite different results, at least for some companies, it is important to gain an insight into the potential impact of these differences. Therefore, OLS regression analyses are performed to assess whether the different methods may, when employed in regression analysis, result in different conclusions being drawn. Four independent variables are employed. Size is the variable most widely used in the literature to explain the variation in voluntary disclosure. Previous studies looking at FL information have generally found size to be important (Clarkson et al., 1994; Frankel et al., 1995; Cahan and Hossain, 1996; Clarkson et al., 1999; Miller and Piotroski, 2000; Johnson et al., 2001; Robb et al., 2001; Walker and Tsalta, 2001; Kent and Ung, 2003; Vanstraelen et al., 2003; Hossain

et al., 2005; Celik et al., 2006; Wasley and Wu, 2006; Lim et al., 2007; Bozzolan and Ipino, 2007; Beretta and Bozzolan, 2008). In contrast to this, three other variables; capital need, earnings volatility and competition rate are also included because prior studies have reached conflicting conclusions. As such, if the five different measures are to yield conflicting results, this is likely to be found when including these variables in the regression analysis.

The need for external funds is likely to exert pressure on companies to provide voluntary information to reduce information asymmetry (Healy and Palepu, 2001). Diamond and Verrecchia (1991) argued that greater disclosure improves stock market liquidity by reducing the cost of capital through either decreased transactions costs or increased marketability of a firm's securities. Previous studies have examined the relationship between the disclosures of FL information and companies' needs for external funds (Clarkson et al., 1994; Frankel et al., 1995; Cahan and Hossain, 1996; Clarkson et al., 1999; Miller and Piotroski, 2000; Johnson et al., 2001; Kent and Ung, 2003; Hossain et al., 2005; Bozzolan and Ipino, 2007). For instance, Clarkson et al. (1994) and Frankel et al. (1995) found that firms are more likely to issue forecasts if they are seeking external finance. Clarkson et al. (1999), Hossain et al. (2005) and Bozzolan and Ipino (2007) demonstrated a positive relationship between forward-looking information and new issues of equity funds. In contrast, Cahan and Hossain (1996), Johnson et al. (2001), and Kent and Ung (2003) found no support for the need for external finance and earnings forecasts.

Earnings volatility is likely to increase information asymmetry (Brown and Hillegesit, 2007) and consequentially increased disclosure may be used to reduce this asymmetry. While a number of previous studies have examined the relationship between FL information and earnings volatility (Miller and Piotroski, 2000; Walker and Tsalta, 2001; Kent and Ung, 2003; Vanstraelen et al., 2003; Wasley and Wu, 2006), the results have been conflicting. For example, Miller and Piotroski (2000) show that firms with more persistent earnings are more likely to provide FL information and Kent and Ung (2003) detect that companies with less volatile earnings are more likely to provide earnings forecast information. Walker and Tsalta (2001) instead document no relationship between earnings volatility and disclosure quality while Wasley and Wu (2006) found a positive relationship between the disclosure of cash flow forecasts and earnings volatility. Zechman (2010) report that managers factor voluntary disclosure into the financing decision to either offset or maintain uninformative reporting. Consistent with information asymmetry theories, prior studies indicate that firms which

frequently disclosed earnings forecast are less likely to experience volatile earnings than firms which infrequently disclose forecasts (Waymire, 1985; Lev and Penman, 1990). On the other hand, Kent and Ung (2003) argued that legal liability occurred from inaccurate forecasts. That is, if the management frequently provides inaccurate FL information, the capital market might discredit any future performance.

Healy and Palepu (2001) argue that companies compete primarily on the basis of price or capacity decisions. Based on this proposition, when companies are competing on capacities, firms try to disclose more information based on economic theory (Shin, 2002). Companies facing capacity (product) competition need funds to finance their capital investment to increase their market share; hence, they are expected to disclose more to reduce information asymmetry and in turn, to reduce the cost of capital. However, previous studies have examined the relationship between FL information and competition (e.g., Kent and Ung, 2003), and found no relationship between competition and FL disclosure.

All financial data was taken from DataStream for the 2005 and 2006 financial years. The natural log of sales was employed as the proxy for size. Capital needs were measured as the change in capital and reserves deflated by year-end total assets, earnings volatility was measured as change in after tax profit deflated by after tax profits of 2005 while competition rate was proxied by total sales of the largest four firms in the industry (based on LSE Industry Classification Benchmark system) divided by total industry sales.

Table 4 reports on the dependent variables while Table 8 reports on the independent variables. The independent variables are tested for multicollinearity. Only one correlation being significant at 5% (Size and competition rate -0.478) and all VIFs being small, hence, correlations nor VIF indicate any problems of multicolinearity.

INSERT TABLE 8

Table 9 presents the results of the multiple regressions for the alternative measures. While all five regressions are significant and all five find that both size and competition are significant some interesting differences emerge. As expected, the weighted and un-weighted disclosure

indices result in nearly identical regressions, however, these are quite different from the results generated by the other three methods. For the two regressions based upon disclosure indices, only size and competition rate are significant. In contrast, when using any of the other three methods, earnings volatility is found to also be significant, while when using the manual content analysis method, capital need is also found to be significant at 10% (but it the opposite direction to what would be expected)[8].

INSERT TABLE 9

Again, these results can be compared to prior studies of corporate social or sustainability disclosures. Here, the results are quite different from those of Joseph and Taplin (2011) and also, to a lesser extent, from those of Haniffa and Cooke (2006). Joseph and Taplin (2011) found a lower R-squared (0.49 vs. 0.57) and fewer significant independent variables when disclosure was measured in terms of length or abundance while Haniffa and Cooke (2006) show similar results for 1996 (0.30 vs 0.48) and only slightly higher R square in 2002 (0.45 vs 0.44) for textual analysis or space devoted to disclosure. This is in contrast to the results found here, when the unweighted index is found to be very much less informative than the manual text unit count in particular (R-square 0.33 vs. 0.68). Joseph and Taplin (2011) argue that their results suggest that sentences are a less accurate measure of disclosure, as they may capture repetition and differences in reporting styles. Now, the R-Square is highest for manual and NVivo textual analysis, while the manual analysis also results in more significant independent variables, suggesting that differences in sentence structures may indeed be a problem. However, these results in contrast suggest that for FL information, there is not a problem of repetition and all of the quantity measures are much more useful than disclosure indices.

6. Discussion of the alternative disclosure methods

We aim, here, to make a comparison among alternative disclosure indices and textual content analysis methods. Comparisons were performed using several sets of analyses. While the results of the correlation reveal a high consistency level, suggesting that the choice of

[8] A possible reason for this is because UK corporate debt is predominately private (Ball et al., 2000). Thus, private debt reduces information asymmetry between managers and lenders

disclosure measure does not matter, further analyses in contrast clearly suggest that the method of measurement is important and different measures may result in quite different conclusions being drawn from any empirical study using disclosure measures as dependent variables.

Similar to previous studies (eg., Chow and Wong-Boren 1987), the correlations, an analysis of the ranks of each company and the regression analysis all suggest that it does not matter greatly if a disclosure index is weighted or un-weighted. However, disclosure indices only record whether or not an item is disclosed, they do not measure the extent of disclosure, so they ignore the amount of space devoted to the topics. The three textual content analysis methods instead capture the volume, in that each instance of disclosure is recorded. While the correlations were all very high across all five measures, suggesting it does not matter which type of measure is employed, in contrast, the T-test, the rankings and the regression analysis clearly suggest that this distinction is important. NVivo and the manual content analysis perform coding at the level of a text unit while Nudist performs coding at the level of a sentence and the rankings suggest that while some companies are not significantly affected by which textual content analysis method is employed others are affected, either because they have a sentence structure that is atypical so that sentences contain more or less text units than average or because the manual analysis typically recognised fewer relevant textual units. While the rankings suggest, that at least for a significant number of companies, it matters which measure is used, the regressions clearly show that the three textual content analysis measures produce a measure of disclosure that is easier to explain, in that there is a far stronger relationship between disclosure and the corporate characteristics that theory and past empirical studies suggest are important determinants of corporate disclosure. To that extent, these three measures all better capture corporate disclosure in a meaningful way. However, these results are in conflict with those of Joseph and Taplin (2011), suggesting that sentences are less useful as they contain repetition. Two possible reasons may account for this difference. Firstly, there are methodological differences with Joseph and Taplin (2011) including charts and tables and coding each bullet point or line as a sentence, while this study followed the more common approach (eg., Walker and Tsalta 2001, Hussainey et al, 2003; Beattie et al, 2004) of excluding all images, charts, tables, pictures and graphs. Secondly, Joseph and Taplin (2011) looked at sustainability disclosures rather than FL information. Thus, their index included items such as a general statement of policy on: the environment; social activities; and, the local economy. They also included items such as waste management

methods, social activities description and stakeholder engagement. All of these are items where disclosure might be very short and succinct or it might be very long and verbose. A long statement might quite easily contain little extra information that will influence users. As such, it might be that sustainability or social information disclosure is more prone to include long statements that have an element of repetition or are not very much more useful than shorter statements. The question of whether this is the case and whether there are systematic differences in how social and environmental disclosures should be measured in contrast to information that is more financially orientated remains a question for further research.

Irrespective of the relative advantages of using textual analysis, the regressions suggest that analysis by text unit is better than the analysis by sentence, supporting the arguments of Unerman (2000) and Beattie et al (2004) that sentences are not the most appropriate volume measure. However, this can be done manually or via a computerised count. Manual textual analysis results in a lower mean score than does computerised textual analysis by text unit, although the R-square is high at 0.8871 and, while it yields a higher R-square, the differences are not enormous. It must also be remembered that manual analysis is more time consuming if large samples are being employed (Hussainey et al., 2003). To this extent, whether or not a manual or a computerised text-unit analyse is optimal will likely depend on the nature of research questions being addressed, on the nature of narratives, and on the sample size. For example, when searching for fairly straightforward subjects such as, for example 'employees' it is easier to capture and measure disclosure, but when the subject is more complex, for example, "social and environmental information", more classification rules are required to measure the disclosures. Establishing relevant computerised coding queries then proves to be much more labour intensive with many iterations being required during which computerised and manual content analysis should be compared and the computerised codes changed to better reflect the results obtained by manual inspections. However, as the sample size increases, automated content analysis offers significant time savings which will more than outweigh the additional time required to set-up the search rules.

7. Summary

The aim of this study is to make comparison among alternative disclosure indices and content analysis methods for the UK non-financial companies before the financial crisis. In the analysis, we employ alternative methods of disclosure indices and content analyses (un-weighted index, a weighted index, manual content analysis, and automated content analysis

using NVivo and Nudist) to investigate the impact of method choice on the measurement of FL information. Comparisons among indices were performed using several sets of analyses. While the results of the correlation reveal that the disclosure scores generated by the alternative methods result in a high level of consistency, further analysis suggests that the method of measurement does matter.

The correlation, t-test, regression and analysis of the ranks of each company suggest that it does matter if a disclosure index is weighted or un-weighted. However, disclosure indices only record whether or not an item is disclosed, they do not measure the extent of disclosure. Content analysis methods instead all capture volume, in that each instance of disclosure is recorded. The results of the regression analysis clearly suggest that this distinction is important and that it is easier to explain the volume of disclosure as measured by content analysis methods.

On average, the results provide quite similar inferences; hence, a trade-off should clearly be made. Based on this, it has been concluded that computerised content analysis using QSR NVivo 8 software, while involving considerable set-up costs, offers a significant time savings when applied to a relatively large sample of narratives. This result is consistent with Beattie et al. (2007) who perform coding by manual content analysis using text unit as am measure of unit since sentence is considered too long to capture intellectual disclosure. Automated content analysis is therefore applied based on an identifying list of FL keywords and a set of topic keywords. Text search queries in QSR NVivo 8 are employed to search for FL keywords and topic keywords, and the results are saved in separate nodes. A special feature in NVivo allows for finding the intersections between the previous nodes using advanced coding queries. The results of intersections are merged in one node, which represents the number of text units of FL information for each company. The overall results are exported to an Excel sheet showing a list of companies' DataStream codes and the number of FL text units for each company.

NVivo and the manual content analysis perform coding at the level of a text unit while Nudist performs coding at the level of a sentence. While often these three methods yielded similar result it is also clear that some companies have a sentence structure that is atypical so that sentences contain more or less text units than average. The regressions suggest that analysis by text unit is the best method and that the choice to be made is simply whether to do this

manually or to use a computerised method. Which should be used appears not to mater in terms of the impact upon the results generated. Rather, the choice will depend on the nature of research questions being addressed, on the nature of narratives, and on the sample size. For example, when the sentence consists of a small number of words it is easier to capture the disclosure score, but when the length of the sentence increases more classification rules are required to measure the disclosures. Establishing relevant computerised coding queries then proves to be much more labour intensive with many iterations being required during which computerised and manual content analysis should be compared and the computerised codes changed to better reflect the results obtained by manual inspections. However, as the sample size or/and classification rules increase, automated content analysis offers significant time savings which will more than outweigh the additional time required to set-up the search rules. On the other hand, the use of manual content analysis provides a significant time saving if there is a small sample size or if it is employed to search only for relatively small range of disclosures.

Moreover, this study focuses on voluntary disclosure of forward-looking information in the annual report narratives, since annual reports are considered to be the best communication media. Despite the fact that the annual reports suffer from a lack of timelines, since many value relevant events are reflected in stock prices as soon as the information reaches the market while their influence on reported earnings often occurs with a time lag. Therefore, the same methodology could be conducted for other forms of communication such as—press releases, conference calls, corporate websites, or interim reports after—identifying a new list of forward-looking keywords which is more relevant to information disclosed in such media forms.

References

Abad, C., Bravo, F. and Trombetta, M. (2008), "Methodological and Empirical Analysis of Problems in the Measurement of Voluntary Disclosure", Workshop on Accounting and Finance, June 19- June 20, University Bocconi.

Abrahamson, E. and Amir, E. (1996), "The Information Content of the President's Letter to Shareholders", *Journal of Business Finance and Accounting*, 23.8: 1156-1182

Accounting Standards Board (ASB) (1993), *Operating and Financial Review*, (London, ASB).

Accounting Standards Board (ASB) (2003), *Operating and Financial Review*, (London, ASB).

Accounting Standards Board (ASB) (2005), *Reporting Statement (RS) 1: Operating and Financial Review*, (London, ASB).

Accounting Standards Board (ASB) (2006), *Reporting Statement (RS): Operating and Financial Review*, (London, ASB).

Adams, C.A., Hill, W.Y. and Roberts, C.B. (1998) "Corporate Social Reporting Practices in Western Europe: Legitimating Corporate Behaviour", *The British Accounting Review*, 30.1: 1-21.

Ahmed, K. and Courtis, J. K. (1999), "Association between Corporate Characteristics and Disclosure Levels in Annual Reports: A Meta-Analysis, *The British Accounting Review*, 31.1: 35-61.

Aljifri, K. and Hussainey, K. (2007), "The Determinants of Forward-Looking Information in Annual Reports of UAE Companies", *Managerial Auditing Journal*, 22.9: 881-894

Al-Najjar B. and Abed S. (2014), "The Association between Disclosure of Forward-looking Information and Corporate Governance Mechanisms: Evidence from the UK before the Financial Crisis Period" *Managerial Auditing Journal*, 29 (7), 578 - 595

Athansakaou, V. and Hussainey, K. (2014), "The perceived credibility of forward-looking performance disclosure, Accounting and Business Research, Forthcoming

Beattie, V., McInnes, B. and Fearnley, S. (2004), "A Methodology for Analysing and Evaluating Narratives in Annual Reports: A Comprehensive Descriptive Profile and Matrices for Disclosure Quality Attributes", *Accounting Forum*, 28: 205-236.

Beattie, V. and Thomson, S. J. (2007), "Lifting the Lid on the Use of Content Analysis to Investigate Intellectual Capital Analysis", *Accounting Forum*, 31: 129-163

Beretta, S. and Bozzolan, S. (2008),"Quality versus Quantity: the Case of Forward-Looking Disclosure", *Journal of Accounting, Auditing and Finance*, 23.3: 333-375.

Botosan, C. A. (1997), "Disclosure Level and the Cost of Equity Capital", *The Accounting Review,* 72.3: 323–350.

Bozzolan, S. (2003) "Italian Annual Intellectual Capital Disclosures", *Journal of Intellectual Capital,* 4.4: 543-558.

Bozzolan, S. and Mazzola, P. (2007), "Strategic Plan Presentations to Financial Analysts: the Effects on Earnings Forecasts' Revision and Cost of Capital", Working Paper.

Bozzolan, S. and Ipino, E. (2007),"Information Asymmetries and IPO Underpricing: The Role of Voluntary Disclosure of Forward-Looking Information in the Prospectus", Working Paper, University of Padova, Italy.

Breton, G. and Taffler, R. J. (2001), "Accounting Information and Analysts Stock Recommendation Decisions: A Content Analysis Approach", *Accounting and Business Research,* 31.2: 91-101.

Brown, S. and Hillegeist, S. L. (2007), "How Disclosure Quality Affects the Level of Information Asymmetry", *Review of Accounting Studies,* 12: 443-477.

Brown, S. and Tucker, J. (2011), "Large-sample evidence on firms' year to year MD&A Modifications", Journal of Accounting Research, 49 (2), 309-346.

Buzby,S.L. (1974) "Selected Items of Information and Their Disclosure in Annual Reports", *Accounting Review,* 49.3: 423-435.

Cahan, S. F., and Hossain, M. (1996),"The Investment Opportunity Set and Disclosure Policy Choice: some Malaysian Evidence", *Asia Pacific Journal of Management,* 13.1: 65-85.

Campbell, D. (2000), "Legitimacy theory or managerial reality construction. Corporate social disclosure in Marks and Spencer corporate reports, 1969-1997", *Accounting Forum,* 24.1:80-100

Campbell, D. (2004), "A longitudinal and cross-sectional analysis of environmental disclosure in UK companies – A research note", *British Accounting Review,* 36.1: 107-118

Celik, O., Ecer, A. and Karabacak, H. (2006), "Disclosure of Forward- Looking Information: Evidence from Listed Companies on Istanbul Stock Exchange (ISE)", *Investment Management and Financial Innovations,* 3.2: 197-216.

Cerf, A. R. (1961), *Corporate Reporting and Investments Decisions,* University of California, Berkeley

Cho, C.H. (2009), "Legitimation Strategies Used in Response to Environmental Disasters: a French Case Study of Total SA's Erika and AZF Incidents", *European Accounting Review,* 18.1: 33-62.

Cho, C.H. and Patten, D.M. (2007), "The Role of Accounting Disclosures as a Tool of Legitimacy: A Research Note", *Accounting, Organizations and Society,* 32.7/8: 639-647.

Cho, C.H, Roberts, R.W. and Patten, D.M. (2010). "The Language of US Corporate Environmental Disclosure", *Accounting, Organizations and Society,* 35: 431-443.

Chow, C. W., and Wong-Boren, A. (1987), "Voluntary Financial Disclosure by Mexican corporations", *Accounting Review*, 62.3: 533-541

Clarkson, P. M., Kao, J. L., and Richardson, G. D. (1994),"The Voluntary Inclusion of Forecasts in the MD & A Section of Annual Reports", *Contemporary Accounting Research*, 11.1: 423-450

Clarkson, P. M., Kao, J. L., and Richardson, G. D. (1999),"Evidence that Management Discussion Analysis (MD&A) is A Part of A Firm's Overall Disclosure Package", *Contemporary Accounting Research,* 16.1: 11-35.

Core, J. (2001), "A Review of the Empirical Disclosure Literature: Discussion", *Journal of Accounting and Economics*, Vol. 31, PP. 441-456.

Coy, D. and Dixon, K. (2004), "The Public Accountability Index: Crafting a Parametric Disclosure Index for Annual Reports", *The British Accounting Review*, 36.1: 79-106.

Davis, A. and Tama-Sweet, I. (2012), "Managers' use of language across alternative disclosure outlet: earnings press release versus MD&A, *Contemporary Accounting Research*, Vol. 29, Issue 3, PP. 804-837.

Deegan, C. (2002), "The Legitimising Effect of Social and Environmental Disclosures: A Theoretical Foundation", *Accounting, Auditing & Accountability Journal,* 15.3: 282-311.

Deegan, C, Rankin, M and Voight, P. (2000), "Firms' disclosure reactions to major social incidents, Australian evidence", *Accounting Forum,* 24.1: 101-130

Deumes, R. (2008),"Corporate Risk Reporting: A Content Analysis of Narrative Risk Disclosure in Prospectuses", *Journal of Business Communication*, 45.2: 120-157.

Diamond, D. W. and Verrecchia, R. E. (1991),"Disclosure, Liquidity, and the Cost of Capital", *The Journal of Finance*, 46: 1325-1360.

DiMaggio, P. and Powell, W. (1991), "Introduction", in Powell, W. and DiMaggio, P. (eds.), *The New Institutionalism in Organizational Analysis*, The University of Chicago Press, London, PP. 1-38.

Elzahar H., Hussainey K., (2012) "Determinants of narrative risk disclosures in UK interim reports", *Journal of Risk Finance*, 13 (2): 133 – 147.

Eng, L.L. and Mak, Y.T. (2003), "Corporate Governance and Voluntary Disclosure", *Journal of Accounting and Public Policy,* 22: 325-345

Ernst and Ernst (1978) *Social Responsibility Disclosure: 1977 Survey,* Ernst & Ernst.

Frankel, R., McNicholls, M. and Wilson, G. P. (1995), "Discretionary Disclosure and External Financing", *The Accounting Review*, 70: 135-150.

Firth, M. (1979), "The Impact of Size, Stock Market Listing and Auditors on Voluntary Disclosure in Corporate Annual Reports", *Accounting and Business Research*, 9.36: 273-280.

Gaver, J. and Gaver, K. (1993), "Additional Evidence on the Association between the Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies", *Journal of Accounting and Economics*, Vol. 16, PP. 127-160.

Gisbert A., Navallas B.(2013), "The association between voluntary disclosure and corporate governance in the presence of severe agency conflicts", *Advances in Accounting*, Vol 29, PP. 286-298.

Gray, R., Kouhy, R. and Lavers, S. (1995), "Corporate Social and Environmental Reporting: A Review of the literature and A Longitudinal study of UK Disclosure", *Accounting, Auditing and Accountability*, 8: 47-77.

Gruning, M. (2011), "Artificial Intelligence Measurement of Disclosure (AIMD)", *European Accounting Review*, 20.3: 485-520.

Guidry, R.P. and Patten, D.M. (2012), "Voluntary Disclosure Theory and Financial Control Variables: An Assessment of Recent Environmental Disclosure Research", *Accounting Forum*, 36: 81-90.

Hackston, D. and Milne, M J. (1996), "Some Determinist of Social and Environmental Disclosure in New Zealand Companies", *Accounting, Auditing and Accountability Journal*, 9.1: 77-108

Haniffa, R.M. and Cooke, T.E. (2005), "The Impact of Culture and Governance on Corporate Social Reporting", *Journal of Accounting and Public Policy*, 24.5: 391-430.

Healy, P. and Palepu, K. (2001), "Information Asymmetry, Corporate Disclosure, and the Capital markets: A Review of the Empirical Disclosure Literature", *Journal of Accounting and Economics*, 31: 405-440.

Holsti, O. R. (1969), *Content Analysis for the Social Sciences and Humanities*, Addison-Wesley, Reading MA.

Hossain, M., Ahmed, K. and Godfrey, J. M. (2005),"Investment Opportunity Set and Voluntary Disclosure of Prospective Information: A Simultaneous Equations Approach", *Journal of Business Finance and Accounting*, 32. 5/6: 871-907.

Hussainey, K. and al-Najjar, B. (2011), "Future Orientated Narrative Reporting: determinants and Use", *Journal of Applied Accounting Research*, 12.2: 123-138.

Hussainey, K., Schleicher, T. and Walker, M. (2003), "Understanding Large-scale Disclosure Studies when AIMR-FAF Ratings are not Available: the Case of Prices Leading Earnings", *Accounting and Business Research*, 33.4: 275-294.

Hussainey, K. and Walker, M., (2009), "The effects of voluntary disclosure and dividend propensity on process leadings earnings", *Accounting and Business Research,* 39.1: 37-55

Hutton, A. P., Miller, G. S. and Skinner, D. J. (2003), "The Role of Supplementary Sentiments with Management Earnings Forecasts", *Journal of Accounting Research,* 41.5: 867-890

Johnson, M. F., Kasznik, R. and Nelson, K. K. (2001), "The Impact of Securities Litigation Reform on the Disclosure of Forward-Looking Information by High Technology Firms", *Journal of Accounting Research,* 39.2: 297-327.

Joseph, C., and R. Taplin. 2011. "The Measurement of Sustainability Disclosure: Abundance versus Occurrence", *Accounting Forum,* 35 (1): 19-31.

Karamanou, I., & Vafeas, N. (2005). The Association between Corporate Boards, Audit Committees and Management Earnings Forecasts: An Empirical Analysis. Journal of Accounting Research, 43(3), 453–485.

Kent, P. and Ung, K. (2003),"Voluntary Disclosure of Forward-Looking Earnings Information in Australia", *Australian Journal of Management,* 28.3: 273-286

Kim, E-H and Lyon, T.M. (2011), "Strategic Environmental Disclosures: Evidence From DOE's Voluntary Greenhouse Gas Registry", *Journal of Environmenatl Economics and Management,* 61.3: 311-326.

Kothari, S. P., Li, X. and Short, J. E. (2009), "The Effect of Disclosure by Management, Analysis, and Business Press on Cost of Capital, Return Volatility, and Analysts Forecasts: A Study Using Content Analysis", *The Accounting Review,* 84.5: 1639-1670.

Krippendorff, K. (2004), *Content Analysis.* Thousand Oaks, CA: Sage.

Lim, S., Matolcsy, Z. and Chow, D. (2007), "The Association between Board Composition and Different Types of Voluntary Disclosure", *European Accounting Review,* 16.3: 555-583

Lyon. T.P. and Maxwell, J.W. (2011), "Greenwash: Corporate Environmental Disclosures Under Threat of Audit", *Journal of Economics and Management Strategy,* 101: 3-41.

Marston, C. L. and Shrives, P. J. (1991), "The Use of Disclosure Indices in Accounting Research: A Review Article", *British Accounting Review,* 23: 195-210.

Meek, G., Roberts, C.B. and Gray, S.J. (1995), "Factors influencing voluntary annual report disclosures by U.S., U.K. and continental European corporations", *Journal of International Business Studies,* 26.3: 555-572

Miller, G. S. and Piotroski, J. D. (2000), "Forward-looking Earnings Statement: Determinants and Market Response", Working paper. Harvard Business School, July.

Milne, M. J. and Adler, R. W. (1999), "Exploring the Reliability of Social and Environmental Disclosures Content Analysis", *Accounting, Auditing and Accountability Journal,* 12.2: 237-256.

O'Sullivan, M., Percy, M. and Stewart, J. (2008), "Australian Evidence on Corporate Governance Attributes and their Association with Forward-Looking Information in the Annual Report", *Journal of Management and Governance*, Vol. 12, PP. 5-35.

Robb, S. W. G., Single, L. E. and Zarzeski, M. T. (2001), "Nonfinancial Disclosures across Anglo-American Countries", *Journal of International Accounting, Auditing and Taxation*, 10: 71-83

Shin, Y-C. (2002), "The Effect of Product Market competition on Corporate Voluntary Disclosure Decisions", Tulane University, A. B. Freeman School of Business, Dec.

Singhvi, S. S. and Desai, H. B. (1971), "An Empirical Analysis of the Quality of Corporate Financial Disclosure", *Accounting Review*, 46.1: 129-138

Smith, C. and Watts, R. (1992), "The Investment Opportunity Set and Corporate Financing and Dividend Policies", *Journal of Financial Economics*, Vol. 32, PP. 263-292.

Unerman, J. (2000), "Methodological Issues Reflections on Quantification in Corporate Social Reporting Content Analysis", *Accounting, Auditing and Accountability Journal*, 13.5: 667-680

Vanstraelen, A., Zarzeki, M. T. and Robb, S. W. G. (2003),"Corporate Nonfinancial Disclosure Analyst Forecast Ability across Three European Countries", *Journal of International Financial Management and Accounting*, 14.3: 249-278

Walker, M. and Tsalta, A. (2001), *Corporate Financial Disclosure and Analyst Forecasting Activity: Preliminary Evidence for the UK*, ACCA Research Report, 67.

Wang M. and Hussainey K. (2013),"Voluntary forward-looking statements driven by corporate governance and their value relevance" *Journal of Accounting and Public Policy*, 32 (3): 26 – 49.

Wasley, C. E. and Wu, J. S. (2006),"Why Do Managers Voluntarily Issue Cash Flow Forecasts?", *Journal of Accounting Research*, 44.2: 389-429.

Zechman, S. (2010), "The relationship between disclosure and financial reporting evidence from synthetic leases", Journal of Accounting Research, 48 (3), 725-765.

Zucker, LG., (1987), "Institutional Theories of Organization", *Annual Review of Sociology*, Vol.13, PP. 443-464

Table 1 List of Forward-Looking Keywords

Number	Keywords	Number	Keywords
1.	Accelerate	19.	Look forward
2.	Anticipate	20.	Next
3,	Await	21.	Near term, medium term
4.	Coming[financial] year[s]	22.	Optimistic
5.	Coming months	23.	Outlook
6.	Confidence, Confident	24.	Plan
7.	Convince	25.	Predict
8.	Future	26.	Remain
9.	Possible	27.	Renew
10.	Estimate	28.	Probable
11.	Aim	29.	Opportunity
12.	Expect	30.	Commitment
13.	Forecast	31.	Further
14.	Forthcoming	32.	Chance
15.	Hope	33.	Well placed, Well positioned
16.	Intend, Intention	34.	Year[s] ahead
17.	Likely, Unlikely	35.	2006/2007, 2006-2007, 2006/07
18.	Look ahead	36.	2007, 2008….2030

Table 2 List of Forward-Looking Topics

Themes	Keywords
Strategy	Mission, vision, strategy, policy, goal, proposal, target, programme, plan, objective
Structure	Expansion, development, modification, improvement, product, invention, growth, progress, challenge, acquisition, merger, take over, market share
Environment	Legal, regulation, law, environment, rule, politics, social, economical, industry, technology, competition, risk, uncertainty, market, trade, demand, inflation, interest rate, service, trend, employee, leadership, oil price, recession, raw material
Financial	Earning, revenue, sales, turnover, cash, debt, loan, leverage, cost, charge, backlog, return, outcome, income, profit, contribution, investment, assets, saving, benefit, dividend, expenditure, expense, payment, tax, liability, obligation, losses, margin, equity, liquidity, fund, depreciation, research and development, ROIC, ROCE, ROE, ROA, EPS, EBIT, FCF, COGS, NPV

Table 3 Example of NUD*IST Command File

(Search-text"[predict]" pattern-search? yes whole-word? yes case-sensitive? no node (2 16) node-title "predict$")	
(	Start command file
Search text	Start a search text

"[	Start accumulate keywords to search
Predict	Search for the keywords predict
]"	End accumulate keywords
Pattern search	No pattern search to find words with similar meanings or with different tenses
Whole word	Search for the whole word
Case sensitive	Do not require case sensitive
Node (216)	Save results in Node 216
Node title	Name node predict$
)	End command file

Table 4 Descriptive Analysis for the Scores Generated by the Alternative Methods

Methods	**Mean**	**Median**	**S.D**	**Minimum**	**Maximum**
Un-weighted (%)	0.37	0.34	0.12	0.19	0.64
Weighted (%)	0.32	0.30	0.11	0.16	0.60
Manual (no units)	54.87	53.00	21.26	25	122
NVivo (no units)	71.83	64.50	37.00	16	197
NUD*IST (no sentances)	48.13	45.50	24.16	15	137

Table 5 Correlations between the Scores Generated by the Alternative Methods

	Weighted	**Manual**	**NVivo**	**NUD*IST**
Un-weighted	.989**	.828**	.810**	.839**
Weighted		.830**	.817**	.847**
Manual			.871**	.842**
NVivo				.922**

** Correlation is significant at the 0.01 level.

Table 6 Mean Comparison T-test

Methods	Mean	Diff	t	Significance
Un-weighted- Weighted	0.37-0.32	0.05	1.682	0.096
Manual-Nvivo	54.8-71.8	-16.96	-2.177	0.035
Manual - Nudist	54.8-48.13	6.74	1.147	0.256
NVivo- Nudist	71.83-48.13	23.7	2.938	0.005

Table 7. Company Disclosure Ranks for each Method

Company Names	Un-weighted	Weighted	Manual	NVivo	Nudist	S.D
AGA	25.00	24.00	21.00	21.00	22.00	1.82
British Airways	28.00	28.00	27.00	28.00	28.00	0.45
British Energy	26.00	27.00	24.00	23.00	26.50	1.72
Britvic	18.50	21.00	15.00	10.50	10.50	4.07
Care UK	11.00	9.00	8.00	12.00	7.00	2.07
Fenner	22.50	22.00	18.00	13.00	18.50	3.82
GKN	30.00	30.00	30.00	30.00	30.00	**0**
Greggs	1.00	3.00	3.00	3.00	4.00	1.10
Hill & Smith	2.00	1.00	1.00	1.00	1.00	0.45
Informa	4.00	7.00	7.00	17.00	15.00	5.66
InterCon.	13.50	15.00	3.00	22.00	20.00	**7.41**
JJB	7.00	6.00	3.00	2.00	2.00	2.35
Kelda	15.50	17.00	20.00	15.00	12.50	2.76
Lookers	3.00	2.00	13.50	8.00	3.00	4.85
M&S	18.50	20.00	25.00	18.00	14.00	3.97
Marston	5.50	4.50	5.00	7.00	7.00	1.15
Melrose Resources	8.50	10.00	9.50	4.00	21.00	6.28
Petrofac	21.00	19.00	19.00	24.00	25.00	2.79
Premier Oil	8.50	8.00	22.50	26.00	16.00	8.10
PZ Cussons	11.00	12.50	9.50	10.50	9.00	1.37
ROK	11.00	12.50	12.00	14.00	10.50	1.37
SCI Ent.	18.50	16.00	11.00	5.00	7.00	5.74
Scot.& Newcastle	29.00	29.00	29.00	25.00	24.00	2.49
Skypharma	22.50	23.00	16.50	20.00	23.00	2.81
Smiths Group	27.00	26.00	28.00	29.00	29.00	1.30
Stagecoach	24.00	25.00	26.00	27.00	26.50	1.20
Uniq	13.50	18.00	13.50	19.00	18.50	2.76
Woolworth	5.50	4.50	6.00	6.00	5.00	0.65
WSP	15.50	11.00	22.50	9.00	17.00	5.30
Yule Catto	18.50	14.00	16.50	16.00	12.50	2.32
Average of Deviations						**2.96**

Table 8. Descriptive Statistics for the independent variables

Methods	Mean	Median	S.D	Minimum	Maximum	VIF
Size	13.378	13.186	1.421	10.33	15.89	1.710
Capital need	.0362	.0380	.0816	-.12	.21	1.349
Earnings volatility	3.756	.1591	16.995	-7.73	83.67	1.334
Competition rate	.4937	.3493	.2686	.21	.99	1.319

Table 9. The Results of Multiple Regression Analyses for the Alternative Methods

Independent Variables	Un-weighted	Weighted	Manual	NVivo	NUD*IST
Intercept	-0.38 -1.436	-0.37 -1.525	**-63.83** **-2.047***	**-161.25** **-2.652****	-50.77 -1.185
Size	**0.05** **2.655****	**0.05** **2.715****	**7.217** **3.285*****	**15.43** **3.601*****	**6.306** **2.089****
Capital need	-0.36 -1.211	-0.33 -1.211	**-63.08** **-1.808***	-33.54 -0.493	9.78 0.204
Earnings volatility	0.002 1.395	0.002 1.587	**0.591** **3.825*****	**1.076** **3.571*****	**0.915** **4.306*****
Competition rate	**0.177** **2.114****	**0.169** **2.220****	**42.19** **4.305****	**51.51** **2.696****	**25.15** **1.868***
Adjusted R^2	0.335	0.367	0.681	0.628	0.580
Model F-Test (P value)	4.029 .015	4.481 .010	13.815 .000	11.109 .000	9.275 .000

*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.