THE ANALYSIS OF INTERIM MANAGEMENT STATEMENT TONE: A COMPARISON OF MANUAL AND AUTOMATED METHODS

Sheehan Rahman[1, 2]

Accounting and Finance Group

Alliance Manchester Business School

The University of Manchester

[1] The author is a final year Doctoral Researcher (viva completed, awaiting degree) at Alliance Manchester Business School, The University of Manchester, UK. For questions and queries please contact the author at: sheehan.rahman@postgrad.mbs.ac.uk

[2] The author thanks Professor Martin Walker, Dr. Thomas Schleicher, and Professor Edward Lee from Alliance Manchester Business School, The University of Manchester and Professor Niamh Brennan from Quinn School of Business, University College Dublin for their valuable guidance and comments for this paper.

Abstract

In this study I contribute to the debate on the usefulness of manual versus automated content analysis methods by comparing the tone score derived from manual content analysis with that from computer assisted content analysis in order to investigate the explanatory power of these two methods. I first discuss the literature on quantifying accounting narratives, describing the advantages and disadvantages of each method. For instance, automated content analysis tools allow processing a large amount of data quickly and clinically. However these are essentially computer-assisted word counts, and are typically limited in capturing the complexities and subtleties involved in disclosure text (Clatworthy and Jones, 2003; Henry, 2008). On the other hand, manual content analysis is time consuming and the coding of texts can be subjective as judgment varies between individual coder. Nevertheless it has been argued that a manual coder can identify the central message of a text in a more reliable way since by nature it is 'meaning-oriented' (e.g. Smith and Taffler, 2000; Clatworthy and Jones, 2003; Schleicher and Walker, 2010). As academics and analysts have to weigh up the costs and benefits of manual and automated methods in decoding tones of financial disclosures, a fundamental question that arises is which of these two methods are better for explaining market reactions to the release of financial disclosures. Thus, I seek to identify which of these two methods yields the better tone score for explaining contemporaneous share price movements.

For this study, I use 1022 IMSs from a random sample of 100 non-financial FTSE All-Share Index firms in the period 2008-2013. I then compute alternative net tone scores by employing (a) manual content analysis to record the tone of statements on financial performance (i.e. positive, neutral or negative) and calculating a net tone score for every IMS based on the number of positive and negative narratives and (b) automated content analysis using a software suite called WEBMATRIX, and using the Henry (2006) List of Positivity and Negativity for calculating a net tone score for every IMS based on the number of positive and negative words. To measure contemporaneous share price movements, I compute the three-day cumulative abnormal returns surrounding the days of IMS release. In order to ensure that my results are not affected by potential endogeneity bias, I supplement ordinary least square (OLS) regressions with fixed-effects (FE) regressions that controls for within-firm variation. I find that the net tone measured by using manual as opposed to automated content analysis technique, has greater explanatory power for share price movements around the time of IMS

disclosure. My results are similar using both forms of regression estimations and for they hold with and without controlling for firm characteristics, industry (or firm) and year effects.

In addition, prior studies suggest that disclosures are often ‘sugar-coated’ with positive messages (Abrahamson and Amir, 1996, p.1163) and that negative messages yield greater market response than positive messages (Tetlock, 2007; Tetlock et al., 2008). This could be due to the use of ‘negated’ positive words in disclosure texts as suggested by Loughran and McDonald (2011) or because of different concealment strategies employed by managers (Merkl-Davies and Brennan, 2007), both of which are less likely to be revealed in computer assisted word counts than in manual content analysis. The strategy of using negated positive tone to frame negative messages is perhaps driven by managers’ belief in prospect theory, a theory of discretionary narrative disclosure, which suggests that investors are likely to respond to information based on the way the message in the information is framed (Henry, 2006). Automated content analysis cannot capture any possible negated words but only counts the number of positive and negative words from a given list. On the contrary, manual content analysis is more likely to code any negated tone with greater objectivity. This motivates me to compare the power of manual measures of positivity and negativity with their automated counterparts for explaining share price movements. I find, using both ordinary least square and fixed-effects regressions that measures of positivity and negativity computed under manual content analysis yield greater explanatory power than the corresponding automated measures of positivity and negativity. This suggests that the greater explanatory power of manual tone score can be extended towards individual measures of positivity and negativity.

Finally, I examine which of these two methods yields a tone that better predicts changes in future firm performance since Davis et al. (2012) suggest that the tone can potentially signal future firm performance. For measuring future firm performance changes, I compute the firm’s impending annual change in (a) operating profit (b) earnings per share (c) return on assets and (d) sales. I find in nearly all cases that the tone derived by manual as opposed to automated content analysis has greater explanatory power for predicting changes in future firm performance. However, I find that the increases in model explanatory power from automated to manual methods is modest in nature both for explaining contemporaneous share price reaction as well as for predicting future firm performance. Therefore academicians have to weigh up the incremental benefits of manual content analysis with its costs, chiefly the time and labour required to process large volumes of financial disclosures.

THE ANALYSIS OF INTERIM MANAGEMENT STATEMENT TONE: A COMPARISON OF MANUAL AND AUTOMATED METHODS

1. Introduction

In recent years an increasing stream of capital market research has developed an interest in how the share price of a firm is influenced by the linguistic tone of accounting narratives, which is broadly viewed as a disclosure's positivity or negativity (Henry and Leone, 2016). Empirical evidence is largely indicative of the linguistic tone of disclosure having incremental information for share prices (Abrahmason and Amir, 1996; Francis et al., 2002; Henry 2006; 2008; Henry and Leone, 2016; Tetlock et al., 2008; Kothari et al., 2009). Although early textual analysis used manual scoring techniques, in recent years computer-assisted word counts have become the norm for computing the linguistic tone of accounting narratives. However, it has often been argued in the accounting domain, without direct comparisons being made, that manual content analysis, though time-consuming and costly, is likely to provide more accurate measures of tone than automated methods (e.g. Clatworthy and Jones, 2003; Schleicher and Walker, 2010). In this study, I compare manual and automated textual analysis methods for computing the tone of accounting narratives, to determine how strongly each is related to stock market reactions, and to determine the net tone score under which method produces the better model for explaining contemporaneous share price movements and predicting future firm performance.

This study has two research questions. The first is to examine if manual content analysis is different from automated content analysis in explaining the market reaction to the tone of financial disclosures. This includes investigating any differences in the explanatory power of the models when tone is split into positive and negative components, since prior studies suggest that the market reaction to negative tone is different from positive tone, perhaps due to scoring errors of negated positive tone in disclosure text (Loughran and McDonald, 2011) as well as due to the differences in credibility for positive and negative narratives to the investor (Henry and Leone, 2016). The second is to examine if manual content analysis is different from automated content analysis in predicting future firm performance.

Manual textual analysis involves human coding of the narrative. In the context of this study, it involves identifying the tone of the narrative. The human coder needs to have an understanding of the business discipline and how particular reported information is likely to affect the firm. A disclosure with a positive (neutral, negative) tone is likely to positively (neutrally, negatively) affect the economic and financial well-being of the disclosing entity. Manual content analysis is meaning-oriented, and therefore it requires the coder to apply a degree of judgment in determining the tone of the narrative, hence it induces some subjectivity in the scoring. It is also time consuming and hence difficult to apply on a large sample of disclosures. However, it can reliably determine the tone if the coder is able to consistently code narratives. Alternatively, automated textual analysis provides the frequency of positive and negative words in a document from positive and negative keyword lists uploaded in a computer programme. The tone can then be calculated as the difference between the frequency of positive and negative words appearing in the document. This form-oriented method of scoring tone is quick and can be applied to a large volume of data but is less reliable if it fails to capture the complexity and subtlety of language that disclosing firms can employ.

The central objective of this study is to compare the net tone score yielded by manual and automated content analysis methods for explaining share price movements and future performance. Interim Management Statements (IMSs) are appropriate financial reports for this kind of comparison between manual and automated tone scores as they are relatively short such that a large volume can be scored by hand and they are required to contain information of the financial performance of the firm. Interim Management Statements are trading statements that firms in EU regulated markets were required to disclose mandatorily from 2007 till 2014. During this period, every firm listed in EU regulated markets were required to publish two IMSs per year, one in the first and one in the third quarter, to report on any material events and transactions during the period and also to describe the financial performance and financial position of the firm, with the disclosing firm retaining the discretion of specific line items that could be reported, and whether to use numbers in their disclosures. Thus, I use 1022 UK Interim Management Statements from a random sample of 100 FTSE All-Share non-financial firms in the period 2008 to 2013. To deter the possibility of my results being affected by endogeneity problems, I supplement ordinary least square (OLS) regressions with fixed-effects (FE) regressions. I find that the tone of an IMS under manual content analysis, as opposed to automated content analysis, has greater explanatory power for share price movements around the time of IMS issuance. My result holds for both ordinary least square and fixed-effects regression models.

Next, prior studies have suggested that negativity measures are more strongly associated with stock market movements than positivity measures. In particular, financial disclosures are ‘sugar-coated’ (Abrahamson and Amir, 1996, p. 1163) and the language in annual report narratives is positively biased (Rutherford, 2005) while there is a stronger relationship between negativity and stock returns than positivity and stock returns (Tetlock, 2007; Tetlock et al., 2008). I find that measures of negativity and positivity under manual content analysis methods yield greater explanatory power for the models of share price movement as opposed to equivalent measures of negativity and positivity in comparable automated methods, and the results hold across ordinary least square and fixe-effects regressions. However, in all cases, the increases in explanatory power from automated methods to manual methods are modest for net tone scores as well as for measures of positivity and negativity. Therefore, the incremental benefits of manual tone in explaining contemporaneous share price movements must be weighed up with the time and costs of manual content analysis.

Finally, Davis et al. (2012) use automated scoring and find a positive association between net tone and future performance. I compare the predictive ability of manual with automated methods for future performance, using four different performance measures—annual changes in operating profit, earnings per share, return on assets and sales. The results differ slightly depending on the type of regression estimations used. Using ordinary least square regressions, I find, in all cases, that the tone under manual method has greater explanatory power for predicting changes in impending annual performance, although the increases in explanatory power from the automated to manual method remain modest to marginal in all cases. Alternatively, using fixed-effects regressions, I find that the manual tone has greater predictive power for signalling changes in operating profit, return on assets and sales when compared to their automated counterparts while the explanatory power in the manual and automated models are virtually the same.

This paper contributes to textual content analysis in the domain of accounting by comparing manual with automated content analysis methods and providing evidence that although the tone computed by manual content analysis, as opposed to automated content analysis, yields greater model explanatory power for contemporaneous share price movements and future firm performance, the differences in explanatory power between the two methods are modest to trivial in all cases, so the costs and benefits of manual content analysis have to be weighed up while choosing the appropriate content analysis method for study.

The remainder of this chapter is organized as follows. Section 2 discusses the benefits and costs of manual and automated content analysis methods and develops the hypotheses. Section 3 discusses the sampling and methodology of tone measurement. Section 4 discusses the results. Section 5 concludes.

2. Quantifying Accounting Narratives Using Alternative Methods

Broadly speaking, content analysis is a 'research technique for the objective, systematic and quantitative description of manifest of content of communications' (Berelson, 1952, p. 18). Although this wide definition does not differentiate between auditory, textual or visual forms of communication, in this study I use the phrase 'content analysis' interchangeably with 'textual analysis'. Textual analysis involves analysing the textual content of a written document. The objective of textual analysis is to determine the central message conveyed by the discloser. Textual analysis of accounting narratives such as in annual reports and earnings press releases have been the subject of interest to researchers for identifying the linguistic tone of the communication. The tone of communication indicates whether a disclosure fundamentally communicates a positive, neutral or negative message about the firm's earnings and economic well-being to investors and analysts. The language used in disclosure can vary across firms, industries and time, but managers typically report financial performance in comparative terms (Davis et al., 2012), and therefore narrative in financial reports can be positive or negative. Prior studies suggest that presenting information in positive tone results in more favourable financial performance evaluations than information in negative tone (Levin et al., 1998). The tone of a narrative, i.e. its positivity or negativity, influences how the disclosed information is understood by the market participants such as investors (Katz, 2001). Textual content analysis can be performed manually or by using a computer assisted programme. Early capital market research assessing the implications of qualitative financial disclosure applied manual content analysis techniques (e.g. Francis et al., 1994; 2002). Manual content analysis involves human reading of the financial disclosure (either the entire document or just specific portions of it) to determine what the information narrated in the disclosure discusses on the financial and economic well-being of the firm, in other words, whether the tone of the communication is positive, negative, or neutral. For instance, if a narrative in an Interim Management Statement reports: 'Revenues in the first quarter are likely to be 10% higher than that of last year.', then the coder should realize that an increase in revenue, a firm fundamental, is likely to positively affect the bottom-line of the firm, and therefore make the firm more lucrative to investors, i.e. enhance its share price. The tone of this narrative is positive. This determination of the tone is ‘meaning-oriented’ and it involves human analysts applying a degree of judgment. Therefore, the quality of the coding is a function of, firstly, their knowledge and understanding of business discipline, secondly, their ability and efficiency in scoring tones objectively and consistently, and thirdly, their specialized knowledge of the disclosing firm, the industry and the economy. Thus, manual content analysis induces a degree of subjectivity, as knowledge and understanding in the areas of business may vary between coders. Needless to say, such a method is time consuming and costly, especially for a researcher who is looking to score the tone of narratives from a large number of financial disclosures. However if the content analyst has sound knowledge of the business discipline, is able to code objectively and consistently, and can understand how particular events and outcomes are likely to affect the firm’s economic well-being and the interests of the users of accounting information, then manual content analysis can accurately determine the central message of communication, i.e. whether the tone of a narrative is positive, neutral, or negative. Further, as part of applying judgment, manual content analysis gives the coder an opportunity to identify what statements / narratives are the most relevant in terms of the describing the financial performance of the disclosing entity, and hence it gives the coder the liberty to decide which narratives to code from the entire document and which narratives to ignore, thereby controlling the cost and time consumption for scoring manually.

Advances in information technology have given rise to specialized computer programmes which can provide the frequency counts of listed words present in a document. As a result, most modern textual analyses of the content use computer analysis (Neuendorf, 2002), which is referred in this study as automated content analysis. Automated content analysis is form-oriented as it treats a document as a bag of words. In essence this method considers a list of selected words present in the document and reports the number of times the words appear (Henry and Leone, 2016). Typically, the researcher needs to have a list of keywords that communicates positive messages, and another list that communicates negative messages. The software simply yields the number of positive and negative keywords in a document from the two lists. Generally, if there are more positive words in a document than negative words, then the document is said to communicate an overall positive message, i.e. the tone of the disclosure is said to be positive. For example, if a document contains five occurrences of words from a positive list (either 5 different positive words appear once, or a single word appearing 5 times, or any other combination), and 3 occurrences of words from a negative list, then the document is said to have a tone score of: (5-3) / (5+3) = 2 / 8 = 0.25. Henry and Leone (2016) argue that after the tone is quantified, researchers can use the scores to examine how the tone affects user decision making. Automated content analysis is useful to a researcher as it can quickly process a large number of documents to determine the number of positive and negative keywords that the document contains. An additional advantage of automated word count is its consistency and lack of subjectivity bias. However automated coding is not reliable for adequately deriving the meaning of the message and experienced providers of financial information may use positive or neutral words in a subtle manner to communicate non-positive messages, to reduce the market penalty of disclosing negative news. This disclosure strategy appears to be grounded in prospect theory, which suggests that investor response to disclosure depends on the way the message is framed. If a negative message is framed with positive words, it is most likely to be viewed more favourably by investors as opposed to the same message framed without positive words (Merkl-Davies and Brennan, 2007).

In addition, computer-assisted word counts fail to differentiate between different meanings of a word as it is not context oriented. For instance, the Harvard Psychological Dictionary classify words such as ‘tax’ and ‘liability’ as negative words although they are not necessarily negative in a financial context, while negatively classified words such as ‘crude’(oil) or ‘mine’ are likely to identify with specific industries rather than explain a negative performance (Loughran and McDonald, 2011). Because of its potential weakness in reliably capturing the central message of financial reports, researchers may need to sample large volumes of data in order to identify trends and patterns while employing automated methods. For example, Henry and Leone (2016) use a sample of over 63,000 earnings announcements for comparing the tone under different automated word lists while Li (2008) uses a sample of over 55,000 annual reports to examine the relationship between readability and earnings using automated content analysis software suites.

While automated content analysis allows us to examine massive volumes of financial disclosures quickly and cheaply, manual content analysis is perceived to be more effective in capturing differences in meaning and context. As an academic or analyst arguably weighs up the relative costs and benefits of these two tone scoring methods to decide which is more appropriate for use, a fundamental issue worth considering is the determination which of these tone scoring methods, manual or automated, is better in explaining share price movements at or around the time information is released to the market. To examine this, I develop the following null hypothesis:

H1: Regressing market returns on manual and automated net tone score yields the same explanatory power for the model.

H1 is a non-directional null hypothesis, and can be rejected on the basis of determining which model yields a greater explanatory power for market returns, i.e. by comparing the Adjusted R-Squared (for OLS regressions) and R-Squared (for FE regressions) between manual and automated models. In particular, Vuong's (1989) test can be used for comparing two regression models. It is typically used for comparing the Adjusted R-Squares between two OLS regressions, but because fixed-effects regression outputs do not provide an Adjusted R-Squared value, I apply it for comparing the R-Squared between two FE regressions. Specifically, the Vuong statistic makes probabilistic statements about the two models and tests which model is closer to the true data generating process, i.e. which regression model best explains random share price movements, in terms of net tone scores.

As a related subject, I find it useful to examine whether the market views positive words as value relevant as positive narratives, and negative words as value relevant as negative narratives. A comparison of manual and automated positivity and negativity measures stem from the comparative value relevance of positive and negative messages to the investors. Prior research suggests positive messages in financial disclosures are not considered as important by the market as negative messages. In a descriptive study, Rutherford (2005) finds that the language in annual report narratives is biased towards the positive. Abrahamson and Amir (1996) focus only on negative statements in their study of the tone of voluntary disclosure and suggest that most positive statements are ritualistic and hence irrelevant. Tetlock (2007) and Tetlock et al. (2008) find a stronger relationship between negativity and stock returns than positivity and stock returns.

There are at least two explanations why negativity measures are considered more important by the market than positivity measures. First, Loughran and McDonald (2011) suggest that positive words in texts are frequently negated, i.e. firms frequently use positive words to frame negative messages whereas negative words are rarely used to convey positive messages. As an example, Loughran and McDonald (2011) suggest that the phrase ‘did not benefit’ frequently appears in disclosures. While the word ‘benefit’ is likely to be picked as positive by an automated scoring method, the phrase conveys a negative message. On the contrary, Loughran and McDonald (2011) suggest that negated negative words, such as ‘not downgraded’, rarely appear in text. The disclosure strategy of negated tone appear also to be grounded in prospect theory, which suggests that the manner in which information is presented by the preparer affects the way it is processed by the user. For instance, Henry (2006) uses prospect theory to suggest that framing financial performance in positive tone causes investors to regard the reported performance in terms of ‘increases’ to reference points, thereby influencing investors’ reaction to disclosure. The issue of negated tone reveals an important distinction between manual and automated scoring methods: automated scoring methods are unlikely to pick up negated positive words but would instead score them as positives in error. Since a manual scoring method is likely to be able to pick up the negative framing in these cases, as would nearly any investor, it provides a primary motivation to compare the negativity and positivity scores under manual and automated content analysis to examine which scoring technique better explains changes in share price movements.

Second, managers have motivations to portray the firm’s financial performance in a favourable light: good performance leads to higher share price, resulting in increased salary, bonus, promotion, greater job security etc. Further, managers have incentives to downplay the negative news in financial disclosures in order to delay the market and managerial penalties of the firm not performing well, i.e. reduction in share price, resulting in pay-cuts, layoffs and punishments (Baginski et al., 2000; Merkl-Davies and Brennan, 2007). Given this, investors and analysts are inclined to believe a negative message more than a positive message because, with the possible exception of stock option issuance, managers typically do not have an incentive to bias the stock price downwards. Merkl-Davies and Brennan (2007) observe that managers use various impression management strategies to present the performance favourably. This includes six concealment strategies: (a) attempts to obfuscate bad news by making texts more difficult to read (‘reading ease manipulation’)[3], (b) using persuasive language (‘rhetorical manipulation’), (c) the attempt to emphasize good news by focusing on

[3] It can be argued that negated tone is merely a form of reading ease manipulation. However, for the purpose of examining the differences in explanatory power to market reaction while comparing negativity and positivity measures, I keep these two arguments separate. Negated tone in texts are likely to be picked up by virtually any investor, but most concealment strategies are likely to be captured only by analysts and sophisticated investors.

positive words, themes, or financial performance (‘thematic manipulation’), (d) visually biasing the manner in which the information is presented (‘visual and structural manipulation’), (e) choosing measures that present current financial performance favourably (‘performance comparisons’) and (f) disclosing one number from several to depict financial performance favourably (‘choice of earnings number’). It can be argued that while automated content analysis is unable to capture most of these subtleties, a manual coder has a better chance to score the tone of narrative objectively for at least some of the concealment strategies. Similarly, at least analysts and sophisticated investors have a chance to capture the meaning conveyed from many of these subtleties. This provides an additional motivation to compare the positivity and negativity scores under manual and automated methods to identify any differences in the explanatory power for share price movements. To examine the relative explanatory power of manual and automated measures of negativity and positivity for share price movements, I develop the following null hypotheses:

H2a: Regressing market returns on manual and automated negativity yields the same explanatory power for the model.
H2b: Regressing market returns on manual and automated positivity yields the same explanatory power for the model.

H2a and H2b are non-directional null hypotheses, and like H1, can be rejected on the basis of identifying which model yields the larger Adjusted R-Squared (for OLS models) or R-Squared (for FE models), i.e. Vuong’s (1989) test can be used to determine which model is closer to the true data generation process for explaining share price movements in terms of positivity and negativity.

The language used in financial disclosures provides managers with opportunities to signal their expectations about future firm performance. Davis et al. (2012) argue that this signalling is done both directly and subtly. This is because the language in disclosures varies significantly across firms, industry and time and ranges from straightforward to promotional (Mahoney and Lewis, 2004). It can be argued that managers typically have incentives to disclose truthfully. For instance, firms with favourable expectations about their future performance would like to disclose them to the market to increase investor’s expectations about earnings. On the contrary, firms with poor expectations about their future performance are also likely to signal them to the market in order to reduce the effect of any share price decline due to negative earnings announcements. Because an Interim Management Statement can be disclosed within a window of ten weeks, it provides managers with an opportunity to signal their expectations about future firm performance prior to a major announcement event such as the interim or annual report. Analysing a large volume of earnings press releases, Davis et al. (2012) find that the net tone score is positively associated with future return on assets (ROA). As I seek to examine the differences between manual and automated methods of measuring tone, it is useful to investigate if the two methods yield the same explanatory power for future performance. I hypothesize:

H3: Regressing future performance on manual and automated net tone score yields the same explanatory power for the model.

H3 is a non-directional hypothesis that can be rejected if the Adjusted R-Squared in OLS regressions or R-Squared in FE regressions of the manual model is different from the corresponding automated model by employing Vuong's (1989) test.

H1 to H3 are all examined by the Vuong (1989) statistic. In supplementary analysis, I put the manual and automated tone measures together in the same model to examine how well they explain variations in share price movements and future firm performance, and then perform tests of difference (T-test) to see if the coefficients of manual and automated tone measures are different.[4]

[4] Overall my arguments for H1 to H3 are implicitly grounded in economics-based theories of discretionary narrative disclosure. In particular, the incomplete revelation hypothesis states that information which are difficult for the user to extract are less reflected in share prices, while information which are easy to extract are impounded in share prices (Merkl-Davies and Brennan, 2007). For instance, H1 and H2 are chiefly concerned with determining which content analysis method has greater explanatory power for contemporaneous share price movements, and I do not specifically emphasize on the magnitude of the tone coefficients in my hypotheses. However, a point to note—the message in negated tone, concealment strategies and subtle use of language are all likely to be more accurately extracted by manual, meaning-oriented content analysis than by automated, computer assisted word-counts. Hence, the share price reactions to manual tone scores as well as the positivity and negativity measures are likely to be stronger than their automated counterparts. This trend can be observed in Tables 3.5 and 3.6, when the magnitude of automated tone measures are found to be smaller when compared to corresponding manual tone measures, a possible reflection of the incomplete revelation hypothesis.

Another economics-based theory of discretionary narrative disclosure pertinent to these hypotheses is the expected utility theory, which suggests that investors may respond to uncertain disclosures based on their perceived informativeness and credibility (Merkl-Davies and Brennan, 2007). H1 and H2 involve regressing market returns to manual and automated net tone scores and their measures of positivity and negativity. Henry (2006) argues that as investors are risk-averse, they react positively to increased disclosure. She suggests that the relation between earnings (i.e. a measure of financial performance, reflected by the tone in this study) and share prices can be strengthened by improving the writing style and verbal content of the disclosure and it is the narrative component in disclosures which contains new information that supplements numerical disclosure.

3. Methodology

3.1 Sample Selection

I aim to obtain a sample of Interim Management Statements that is small enough for handling manual content analysis and at the same time large enough for automated content analysis to adequately identify any relevant trends and patterns that exists. IMSs are appropriate documents for my study as they are (a) generally not longer than two pages in length (and often one page) so manually reading and scoring the entire document can be done fairly quickly and easily and (b) by definition needs to provide at least a general description of the financial performance of the firm so it is suitable for examining the tone of financial performance.

I allocate a financial year to the calendar year in which the majority of months falls. This means if a firm's financial year-end is 31 July, for the purpose of recording, I define its financial year period of 1 August 2007 to 31 July 2008 as the calendar year 2008. Financial years with a June year-end are allocated to the calendar year in which the year-end falls. Since disclosure of IMSs was mandatory for firms having financial year ends on or after 20 January 2007 until November 2014, and the first year of IMS issuance had lower rates of compliance with Disclosure and Transparency Rules (DTR) and 'teething problems' (Schleicher and Walker, 2015), I decide to eliminate 2007 for my study and use a six-year sampling period of 2008 to 2013.

I use 30 June 2008 as the date for sampling, which had 668 firms in the FTSE All-Share Index. I eliminate financial firms as by virtue of the nature of business operation their information on measures of financial performance such as 'sales', 'earnings', 'costs', 'trading' and 'results' is different than that of non-financial firms. I also eliminate firms publishing full quarterly reports as Article 6 of the EU Transparency Directive indicates such firms do not need to disclose an IMS (Deloitte and Touche, 2007). This leaves 324 non-financial IMS disclosing FTSE All-Share Index firms as at 30 June 2008. From this set of firms I randomly select 100 firms for sampling. I notice that out of these 100 firms 15 are FTSE100 firms, 38 are FTSE250 firms, and the remaining 47 are FTSE Small Cap firms, which gives an even representation of these indexes from the FTSE All-Share Index population.

I obtain the Interim Management Statements from Perfect Information Navigator, a corporate information database that has regulatory and non-regulatory news and filings from over 50,000 firms. If all 100 firms in my sample disclosed one IMS in the first quarter and one IMS in the third quarter in each of their financial years during the sampling period, it would result in 1200 IMSs. However, I lose 69 IMSs due to death or delisting and another 109 IMSs as they are either not disclosed by the company and / or are missing from the Perfect Information Navigator. This leaves a final 1022 IMSs for manual and automated content analysis scoring. Table 1 summarizes my sample selection.

[Insert Table 1]

3.2 Tone Measurement

3.2.1 Measuring Tone in Manual Content Analysis

I conduct manual content analysis on the selected sample by reading each Interim Management Statement and identifying statements that describes the financial performance of the firm, whether by use of numbers or in qualitative terms. A statement of financial performance is typically an accounting narrative on financial performance though often a statement is limited to a single sentence. Sometimes a performance statement may encompass more than one sentence, if it captures one particular piece of information. Performance statements on earnings are often believed to have the greatest value relevance. However I do not only score statements related to earnings (or profit) as statements of financial performance. Instead, to identify what constitutes financial performance, as a rule of thumb, I widen the spectrum to encompass six broad topics, namely ‘sales’, ‘earnings’, ‘costs’, ‘trading’, ‘results’ and ‘general unspecified statements (of financial performance)’ or ‘GUS’.[5] Measures of financial performance that fall under any of these categories are scored as statements of financial performance. In other words, statements pertaining to turnover, order book and revenue are classified as ‘sales’ statements; statements on profit or earnings before

[5] This broad definition of what topics constitute statements of financial performance would also assist in making better comparison with my automated content analysis scores (explained in Section 3.2.2), which by default codes the entire IMS document. More importantly, managers are free to choose which line item to report for describing the financial performance in an IMS. Hence, a range of topics are needed for defining financial performance as profits or earnings do not appear in all IMSs. For instance, Schleicher and Walker (2015) find that only 37% (26%) of IMSs report backward- (forward-) looking earnings for the Group or company.

interest and tax (EBIT) are classified as ‘earnings’ statements, and statements describing outlook, success, progress or failure of the firm, or any other narrative are classified as ‘GUS’.

After identifying a performance statement, I code the tone of the narrative—whether it is positive, neutral or negative. Every narrative identified is categorized as consisting of either one of positive, neutral or negative tone. This is done without the aid of any keyword lists; hence it requires me to exercise judgement on how the narrative is likely to affect the financial and / or economic well-being of the firm. As a rule, ‘Positive’ statements are those which have clear or direct indications of improvement or progress from the previous circumstance (e.g. Profit was higher by 10% compared to the first quarter of the corresponding period). These are essentially what one might call good news. ‘Negative’ statements comprise of statements which are clear or direct indications of ‘deterioration’ from the previous position, (e.g. Our Group revenue is expected to be lower in the following quarter due to adverse developments in economy). These are typically what we might call bad news. ‘Neutral’ statements are those which represent the following characteristics: (a) they are neither distinctly positive nor negative, (b) when performance is in line with expectations, (c) when the status quo is preserved (e.g. We began trading modestly in the third quarter, in line with our expectations announced during the half-yearly results). The sample selection illustrated in Table 1 includes the number of positive, neutral and negative statements scored in manual content analysis. Table 2 includes some examples of Positive, Negative and Neutral statements of financial performance.

[Insert Table 2]

I generally look for performance statements that represent the whole company or Group, if available, and do not include statements that are exclusive to a business, geographic or product segment, especially if that statement is also inconsistent with the performance of the Group or company on the same topic. The ability to distinguish group statements from segmental statements is an important advantage of manual content analysis over automated content analysis. Scoring only group statements allows me to capture the overall tone of performance—the central message conveyed in the IMS. For instance, if I find a positive statement on the whole Group on ‘trading’, I code it accordingly. Subsequently, I ignore any negative statements on ‘trading’ that pertains to a specific business, product or geographic segment unless the IMS also contains a negative Group statement on ‘trading’.

After coding the tone of the narrative, I calculate the manual net tone score of each Interim Management Statement. I begin by scoring the values of positivity and negativity to determine how the market would react to positive and negative tone narratives. This is done by following Schleicher and Walker (2010). In the scoring spreadsheet, each statement with a positive (negative, neutral) tone is coded as an indicator variable which takes the value of 1 if the statement is positive (negative, neutral), and zero otherwise. I then compute the manual positivity (negativity) of an IMS, POS-M (NEG-M), by dividing the total number of positive (negative) statements in the IMS with the sum of all positive, negative and neutral statements scored in the IMS. This is shown as follows:

POS-M = POSITIVE-M / (POSITIVE-M + NEUTRAL-M + NEGATIVE-M)
NEG-M = NEGATIVE-M / (POSITIVE-M + NEUTRAL-M + NEGATIVE-M)

Where POSITIVE-M, NEUTRAL-M and NEGATIVE-M are indicator variables taking the value of 1 if the statement coded is positive, neutral and negative respectively, and zero otherwise.

The manual measure of positivity (negativity) computed in the above manner indicates the proportion of positive (negative) statements out of all performance statements coded in an IMS. POS-M and NEG-M range from 0 to 1 and the inclusion of NEUTRAL-M in the denominator makes it directly comparable to automated measures of positivity and negativity (discussed in Section 3.2.2) and also eliminates any potential problems of linear dependency if these measures are used together in a regression model.

Next, I compute the manual net tone score for each IMS as follows:

TONE-M = (POSITIVE-M – NEGATIVE-M) / (POSITIVE-M + NEGATIVE-M)

By definition, TONE-M is a continuous variable that ranges from total negative (-1) to total positive (1). If there are more negative narratives than positive narratives, then the net tone score would range between -1 and 0, and would indicate that the IMS contains more bad than good news. Alternatively, if there are more positive narratives than negative narratives, then the net tone score would range between 0 and 1, and would indicate that the IMS contains more good news than bad news. Absence of any negative (positive) statements would make the tone 1 (-1). A tone score of zero can be achieved if the number of positive or negative statements in the IMS is equal. A step-by-step guideline for manual coding is given in Appendix 1.

3.2.2 Measuring Tone in Automated Content Analysis

For automated content analysis, I use a computer software suite named WEBMATRIX. WEBMATRIX is a word-count software that allows uploading of customized word lists to provide the number (frequency) of words in the document screened that matches a word in that word list. It also provides the total number of words in the IMS document. I use the lists of Positivity and Negativity from Henry (2006) (word lists included in Appendix 2) as the measure of positive and negative tone. The Henry (2006) List is a specialized list of keywords in the accounting domain for automated scoring of positive and negative tone, and according to the findings of Henry and Leone (2016), the Henry (2006) List has greater explanatory power for measuring tone in disclosure research as opposed to non-domain specific generalized wordlists. I upload all IMSs in WEBMATRIX with the Henry (2006) List of Positivity and Negativity to obtain the number of positive and negative words from that list in each IMS document. Then for each IMS I calculate the automated net tone score as follows:

TONE-A = (POSITIVE-A – NEGATIVE-A) / (POSITIVE-A + NEGATIVE-A)

Where POSITIVE-A and NEGATIVE-A refer to the word count frequency from WEBMATRIX based on the positive and negative words in Henry (2006) List.

The automated net tone score, TONE-A is a continuous variable that ranges from total negative (-1) to total positive (1). If there are more negative (positive) words than positive (negative) words from the Henry (2006) List in the IMS document, then the net tone score would range between -1 and 0 (0 and 1), and would indicate that the IMS contains more bad (good) than good (bad) news. The absence of any negative (positive) words would make the tone 1 (-1). A tone score of zero can be achieved if the number of positive or negative words in the IMS is equal. The variable TONE-A is directly comparable to TONE-M since they have the same range (-1, 1). In both these measures, a tone score of 1 indicates a perfectly positive IMS, a tone score between 1 and 0 indicates a predominantly positive IMS, a tone score of 0 indicates a perfectly neutral IMS, one between 0 and -1 indicates a predominantly negative IMS, and a tone score of -1 indicates a perfectly negative IMS.

I compute the positivity and negativity scores as follows:
POS-A = POSITIVE-A / TOTAL where POSITIVE-A is the word count frequency of the IMS document based on the Henry (2006) List of Positivity and TOTAL is the total number of words in the IMS document.

NEG-A = NEGATIVE-A / TOTAL where NEGATIVE-A is the word count frequency of the IMS document based on the Henry (2006) List of Negativity and TOTAL is the total number of words in the IMS document.

In effect the positivity and negativity scores are the percentages of positive and negative words in the whole IMS document. It is worth noting that both manual and automated scores of negativity and positivity have the same range of possible values (0 to 1), and therefore the coefficients of manual and automated measures of positivity and negativity can be directly compared.

In this study, I aim to compare the net tone of IMS computed under manual and automated content analysis methods. The corpus of texts used in manual content analysis in this study is only a subset of the corpus of text for automated content analysis. This is because I compute the automated tone by processing the entire IMS document with WEBMATRIX while I compute the manual tone by selectively choosing statements of financial performance that pertains to the ‘Group’, i.e. the whole firm, and not to any segment, in order to exploit the advantage of manual content analysis. This allows me, the coder in manual content analysis, to pick and choose statements that are likely to be most relevant in determining the tone of financial performance. Therefore, a potential criticism in my research design is whether the tone computed under manual method is sufficiently comparable to the tone computed under automated method. Even though only a subset of all IMS statements are actually scored for manual content analysis, as the coder, I had to read the entire IMS document and decide which statements to score and which statements to ignore, and therefore, the ‘intellectual input’ for the statements not selected for scoring the tone is deemed to be existent even though they are not present in the scoring spreadsheet. Every statement ignored for manual content analysis can be considered as a de-facto 'Neutral' statement, since it has been decided during the manual coding process that they are unlikely to affect the net tone of IMS, i.e. its positivity or negativity. On the other hand, even though there are no lists of 'Neutral' words in Henry (2006), all words outside the lists of Positivity and Negativity are considered to be de-facto 'Neutral' in tone. Given that the formula for computing manual net tone score do not account for neutral statements either in the numerator or in the denominator, I believe the net tone under manual content analysis is not affected in any way by the exclusion of some statements. The same is applicable for automated net tone scores, which do not include neutral words in the numerator or denominator. Therefore, at the very least, I believe that tests for H1 and H3, which involve comparing the net tone scores of manual and automated methods, should have no comparability problems.

For H2, I compare the positivity and negativity scores of manual with automated methods. In this case, the number of neutral statements and words actually scored are part of the denominator for the respective manual and automated formulas for positivity and negativity. Here, I follow the approach of Henry and Leone (2009) and include the total number of statements or words in the denominator for positivity and negativity of each method, in order to avoid linear dependency of these coefficients in the regression models. The total word count in an IMS is automatically derived from the WEBMATRIX output. All the words outside the Negativity and Positivity lists of Henry (2006) are considered to be neutral words and are form part of the denominator (which includes the total IMS word count) in automated negativity and positivity. In contrast, only the number of 'Neutral' financial performance statements actually scored manually is part of the calculation of manual negativity and positivity measures. Hence for the tests in H2, potentially including any 'de-facto' neutral statements (that may have been ignored during manual content analysis) in the denominator of manual positivity and negativity measures are likely to make the coefficients of manual positivity and negativity measures more comparable to their automated counterparts. I address this potential disparity by taking the natural logarithms of manual and automated positivity and negativity measures, which is discussed in detail Section 4.2.2.[6] Using log

[6] An alternative solution to ensure comparability would have been to take a smaller sample of IMSs and then to select and separate the statements scored manually, and then to run the selected statements in WEBMATRIX for obtaining positivity and negativity score for automated method. This can now be readily compared to their manual counterparts, given that the corpus of text is identical. I consider but abandon this idea because it is likely to increase the costs and time needed for automated content analysis substantially.

transformations to enable comparability between alternative positivity and negativity measures is consistent with Henry and Leone (2009).

4. Results

4.1 Descriptive Statistics

Table 3 presents the descriptive statistics for 1022 IMS observations. The mean and median net tone score from automated scoring (mean=0.59 median=0.63) are both higher than that of manual scoring (mean=0.45 median=0.50). A t-test for the difference in means yields p=0.00 and a Wilcoxon's rank sum test for the difference in medians yields p=0.00. Both these net tone scores have a maximum of 1 and minimum of -1, the maximum and minimum possible values within their range respectively. The positive sign of the net tone score under both methods suggests that the average tone of IMS documents is positive. This is consistent with Rutherford's (2005) assertion that financial disclosures contain a greater proportion of positive messages. The positivity (negativity) in automated scoring suggests that a mean of 2.8% (0.67%) and a median of 2.75% (0.5%) of the words in an IMS document match the Henry (2006) List of Positivity (Negativity). This is consistent with prior literature (Abrahamson and Amir, 1996; Tetlock, 2007; Tetlock et al., 2008), and as Rutherford (2005) suggests, this explains why the value of net tone scores are greater than zero.[7] On the contrary, a mean (median) of 62% (67%) of narratives scored in manual content analysis are positive while 22% (20%) are negative. In un-tabulated results on manual scoring, I find that an IMS contains a mean of 2.8 (1.1) statements and a median of 3 (1) statements with a positive (negative) tone on financial performance. This is consistent with the findings of Abrahamson and Amir (1996, p. 1163) that financial disclosures are likely to be 'sugar-coated' with positive statements. While comparing the number of positive and negative words and narratives in an IMS document it is useful to keep an eye on the length of the

[7] The mean and median scores of automated net tone and its positivity and negativity measures can be compared to Henry and Leone (2009) which uses the same lists of positivity and negativity measures. Henry and Leone (2009) employ a US sample of 15526 earnings press releases between 2004 and 2006 and obtains a mean net tone score of 0.44, median net tone score of 0.48, and a mean and median of 0.02 and 0.02 for positivity and 0.01 and 0.01 for negativity respectively. Henry and Leone (2016) employ a US sample of 63357 earnings announcements and obtained a mean net tone of 0.40 and a median of 0.44. I obtain a higher mean and median net tone score than Henry and Leone (2009; 2016) due to a slightly greater number of positive words and marginally lower number of negative words in my sample as opposed to theirs. This might to attributed to (a) a greater number of lawsuits are filed in US as opposed to UK causing US firms to be less optimistic and more cautious in their wording as opposed to UK firms and (b) the difference between an earnings press release or earnings announcements and an Interim Management Statement as the latter only requires to provide a qualitative description of financial performance of the firm and the managers can choose to provide and discuss any line items even without explaining earnings, thereby allowing managers greater freedom to explain and report opportunistically.

whole document. The average length of an IMS is 1008 words as opposed to the median length of 771, reflecting that some companies publish very long documents. Typically these are manufacturing firms that wish to include their quarterly production report in Interim Management Statements.

[Insert Table 3]

I select the following variables for examining their correlations with tone measures:

(a) Cumulative Abnormal Return (CAR): The 3-day (-1, 0, 1) CAR is used as a measure of stock price movements around the date of IMS publication. For abnormal returns I calculate daily market model adjusted returns, u_{it}, as $u_{it} = R_{it} - (\alpha_i + \beta_i R_{mt})$, where R_{it} is the return of firm *i* on day *t*, R_{mt} is the return of the FTSE All-Share Index on day *t* and where R_{it} and R_{mt} are calculated from DataStream Return Indices, RI. α_i and β_i are firm *i*'s estimated market model parameters calculated from the non-event period which runs from day t-60 to day t-10 and from day t+10 to day t+60 relative to the IMS announcement day t=0. CAR is calculated as the sum of the daily market model adjusted returns, u_{it}, over the three-day event period (days t-1, t, t+1), such that $CAR_{it} = u_{it-1} + u_{it} + u_{it+1}$. In Table 3 I report that the median CAR is greater than the mean, suggesting that a larger number of IMSs generate an upward share price reaction, but the magnitude of downward share price reactions are typically greater than upward reactions, consistent with Tetlock et al. (2008).

(b) Firm Size (SIZE): A well-known stock market anomaly is that smaller firms outperform larger firms as they are able to grow at a faster rate, which is reflected in the share price (Simpson, 2014; Ross, 2014). Given this assumption, it is likely that size will have a negative relationship with CAR.[8] So for regressing CAR on tone, I control for firm size over and above the tone of the narrative. SIZE is calculated as the natural logarithm of the market value of equity at the beginning of the year t, calculated as the number of shares multiplied by share price, both at the start of the year t.[9]

[8] However this predicted relationship may not hold if small firms, as opposed to large firms, are not in high growth industries, have lower profitability, or have cash flow stagnation (Israel and Moskowitz, 2013; Ross, 2014).

[9] Another variable often used for examining market response is Analyst Following, which measures the extent to which private information search occurs, and consequently, the extent to which managers produce information to meet the needs of private information search (Baginski et al., 2004). DeFond et al. (2007) find that Analyst Following increases with market response. Following Baginski et al. (2004), I do not include Analyst Following in my study because it is strongly positively correlated with SIZE (r=0.70, p=0.00), suggesting multicollinearity

(c) Profitability Status (LOSS): Investors react differently to information of loss firms as opposed to that of profit firms (Hayn, 1995). So I control for the profitability of the firm by using a LOSS indicator which takes the value of 1 if pre-exceptional operating profit<0 at the start of the year t, and zero otherwise.

(d) Risk (BM): Risk is the uncertainty regarding cash flows in and out of the firm from future projects. The greater the risk of a stock, the greater is the expected return. Firms with extremely high book-to-market ratios are often at risk of financial distress, and larger returns generated by high book-to-market value firms can be a compensation for risk (Galagedera, 2007). Following Link (2012), I compute the ratio of the book-to-market value of the equity at the start of the year t to control for the risk of the firm.

(e) Industry Dummy Variables (INDUSTRY): I compute dummy variables for the firm's industry, to control for the effect of industry differences in a regression of CAR on tone. Following the Industry Classification Benchmark (ICB) for FTSE All-Share Index firms, I create 1/0 indicator variables for each industry. BASICMATERIALS, CONSUMERGOODS, CONSUMERSERVICES, OILANDGAS, UTILITIES, TELECOMMUNICATIONS, TECHNOLOGY and HEALTHCARE are respective indicator variables for the industry of the firm disclosing IMS, each denoting 1 if the firm is classified in the relevant industry and zero otherwise. The dummy variable INDUSTRIALS is omitted. Industry dummy variables are only used in OLS regressions, I substitute them with firm dummy variables for FE regressions to control for any within-firm variation.

(f) Time Effect (YEAR): I compute dummy variables for each year in the sampling period, to control for the effect of time differences in a regression of CAR on tone. YEAR2009, YEAR2010, YEAR2011, YEAR2012 and YEAR2013 are indicator variables to account for the year effect, taking the value of 1 if the IMS was disclosed in the relevant year and zero otherwise. The dummy variable YEAR2008 is omitted.

---

in regressions models if the two variables are used together. However in un-tabulated results, I observe that Analyst Following has expected relationships with all variables, including positive correlations with CAR and positive correlations with both manual and automated measures of positivity and negativity.

An important research design question is the inclusion of unexpected earnings as a measure of control for information in reported net income, often computed as the difference between actual and forecasted EPS, scaled by share price. Unexpected earnings is typically used as a control for regressing CAR on the tone of financial disclosures while examining share price reaction to tone. However, unlike Doran et al. (2012), Davis et al. (2012) and Henry and Leone (2016) all of whom examine earnings press releases, I do not include unexpected earnings as a control variable in my models since Interim Management Statements are not earnings announcements and do not contain income statements: managers have considerable discretion over which line item to report, including whether to include numbers while describing them, and hence it is not necessary to control for the news outside the narratives in IMS. Including a variable for unexpected earnings in my case is likely to distort the coefficient of the tone of the narratives.[10]

I select some additional control variables for examining the impact of tone on the relationship between CAR and future firm performance, based on prior studies that suggest these variables have predictive ability for impending annual changes in measures of firm performance (e.g. Collins et al., 1994; Gelb and Zarowin 2002; Lopes and Walker, 2012). These include: (a) Earnings Yield (EP), defined as operating profit divided by market value of equity at the start of year t, (b) Asset Growth (AG), which is the percentage change in total assets at the start of year t, and (c) Return (RET), which is the financial year buy-and-hold raw returns for the year t. All these variables measure growth in assets and share price, are used as controls in the models regressing changes in future performance on tone alongside SIZE, BM and the industry and year effects.

I measure changes in future firm performance by using four different dependent variables in alternative models with the same regressors: (a) CH_OP, which is the change in pre-exceptional operating profits in year t from year t-1 deflated by total assets at the start of year t, (b) CH_ROA, which is the change in return on assets in year t from year t-1, where return on assets is defined as operating profits divided by total assets (c) CH_EPS, which is the change in earnings per share in year t from year t-1 deflated by assets per share at the start of year t, and (d) CH_SALES, which is the change in sales in year t from year t-1, deflated by

[10] To illustrate, Schleicher and Walker (2015) find that only 4% of forward-looking earnings and 20% of backward-looking earnings information reported in Interim Management Statements is quantitative in nature, and is typically embedded within the narrative.

total assets at the start of year t. These variables are all variations of earnings, profits and sales and indicate future firm fundamentals.

Table 4 presents the Spearman's rank correlations as some variables are discrete while others are continuous. Panel A presents the correlations between the net tone score and measures of tone positivity and negativity under both methods with the variables involved in examining the effect of market reaction on tone: CAR, SIZE, and BM. The variable LENGTH, defined as the number of words in an IMS document, is also included in Panel A for descriptive purposes. CAR, the measure of share price changes, has a greater positive correlation with TONE-M (r=0.19 p=0.00) than with a TONE-A (r=0.12 p=0.00). This provides some prima facie evidence that manual tone could be more closely associated with abnormal market reactions than automated tone. CAR is also positively (negatively) associated with manual and automated positivity (negativity) measures, and while these correlations are all significant at 1%, the correlations are stronger with manual positivity and negativity measures, suggesting that manual content analysis is better in picking up the good and bad news messages conveyed in an IMS. The absolute magnitudes of CAR's correlations with negativity measures are greater than positivity measures for automated scoring. In addition, for manual scoring the absolute magnitude for positivity and negativity are equal. This implies that the market response is more strongly associated with negative words as opposed to positive words, consistent with Tetlock (2007). The market response is almost equally strong for positive and negative narratives, perhaps revealing the effect of negated positive words as proposed by Loughran and McDonald (2011). In addition, CAR is more strongly and negatively (positively) associated with the negativity (positivity) measures of manual scoring than with automated scoring. This provides preliminary evidence that the positivity and negativity of manual scoring has greater explanatory power for market returns than corresponding measures of automated scoring.

The tone scores all exhibit expected associations. For instance, TONE-M is positively associated with TONE-A (r=0.44 p=0.00) but the coefficients have only moderately strong correlation. TONE-M is positively (negatively) associated with both manual and automated measures of positivity (negativity). The automated tone score is also positively (negatively) associated with manual and automated measures of positivity (negativity). These correlations are suggestive of a general trend of consistency between manual and automated scoring methods in terms of intra-tone correlations. Larger firms (SIZE) disclose longer IMSs (with

LENGTH r=0.33 p=0.00) probably due to the scale of their business operations and are less likely to incur a loss (with LOSS r=-0.21 p=0.00). Since I find that larger firms are more likely to be profitable, it is unlikely that the firm size anomaly of a negative relationship with share price will be captured in the market reaction tests. Loss firms (LOSS) disclose longer IMSs (with LENGTH r=0.17 p=0.00) and have positive (negative) correlations with negativity (positivity) measures for automated scores. Loss firms have a negative association with manual negativity, as expected.

Panel B presents the correlations between net tone scores under both methods and variables measuring future firm performance. I find that both TONE-M and TONE-A have a positive association with CH_OP, CH_SALES and CH_EPS. This provides some prima facie evidence that the tone can signal future firm performance. I find that CH_SALES and CH_EPS have a positive correlation with size, earnings yield and annual buy-and-hold raw return, and that CH_OP and CH_ROA have negative correlations with size but positive correlations with BM. As expected, I find that the financial year buy-and-hold raw return RET has positive correlations with both net tone measures (with TONE-A r=0.19 p=0.00 and with TONE-M r=0.19 p=0.00) suggesting that the direction of share price movement over the year is consistent with the net tone of IMSs.

[Insert Table 4]

4.2 Examining Market Reactions to the Tone of Accounting Narratives

4.2.1 Market Reaction to Manual and Automated Net Tone Scores

In the first part of my empirical analysis, I examine market reaction to the tone of accounting narratives. For this I employ a short-window event study to investigate the relation between market reaction and the tone of financial performance, measured alternatively using manual and automated scoring methods. I begin with the following set of ordinary least square (OLS) regressions:

$$CAR = \alpha + \beta_1 TONE\text{-}A + \varepsilon \quad \ldots (1a)$$

$$CAR = \alpha + \beta_1 TONE\text{-}M + \varepsilon \quad \ldots (1b)$$

$$CAR = \alpha + \beta_1 TONE\text{-}A + \beta_2 TONE\text{-}M + \varepsilon \quad \ldots (1c)$$

$$CAR = \alpha + \beta_1 TONE\text{-}A + CONTROLS + \varepsilon \quad \ldots (1d)$$

$$CAR = \alpha + \beta_1 TONE\text{-}M + CONTROLS + \varepsilon \qquad \dots (1e)$$

$$CAR = \alpha + \beta_1 TONE\text{-}A + \beta_2 TONE\text{-}M + CONTROLS + \varepsilon \qquad \dots (1f)$$

The control variables included in eq. 1(d) – 1(f) are firm characteristic variables firm size (SIZE), book-to-market value of equity (BM) and loss indicator (LOSS), industry dummy variables BASICMATERIALS, CONSUMERGOODS, CONSUMERSERVICES, TELECOMMUNICATIONS, TECHNOLOGY, OILANDGAS, UTILITIES, and HEALTHCARE, and time dummy variables YEAR2009, YEAR2010, YEAR2011, YEAR2012 and YEAR2013. All these variables are defined in Section 4.1.

A potential criticism of the models in eq. 1(d) – 1(f) is that they do not take into account the possibility of endogeneity problems arising from omitted variable bias or the tone variable being correlated with the model error term. Since it is not possible to explicitly control for all unobservable factors in OLS regression models, I cannot rule out the possibility of endogeneity bias. This kind of problem can be addressed by using fixed-effect (FE) regressions. The objective of fixed-effects regression for my estimates is to hold constant the average tone effects for each firm. It computes the *mean* and *variation within the mean* of the change in net tone scores for each firm across all the years in my sample, and then it regresses the change in operating profit on net tone score and a dummy variable for every firm in the model that controls for variation 'within-firm'. One firm-dummy has to be omitted from the regression (Clark and Linzer, 2015). I devise the following fixed-effects regression models to supplement eq. 1(d) – 1(f):

$$CAR = \beta_1 TONE\text{-}A + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \qquad \dots (1g)$$

$$CAR = \beta_1 TONE\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \qquad \dots (1h)$$

$$CAR = \beta_1 TONE\text{-}A + \beta_2 TONE\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \qquad \dots (1i)$$

The control variables in the FE regression models in eq. 1(g) – 1(i) differ from their OLS counterparts in that they do not contain the 1/0 indicator variables for industry. This is due to the existence of firm fixed-effect dummy variables that account for within-firm variation, eliminating the need to control for differences in industry.

The results of eq. 1(a) – 1(i), presented in Table 5, suggest that models with the manual net tone score have greater explanatory power for share price movements as opposed to models with automated net tone score. This is evident with and without the inclusion of control variables, and the results hold for both ordinary least square and fixed-effects regression estimates. However, the increases in explanatory power from the automated to manual method, in all cases, is modest in magnitude. For instance, for OLS regressions, the Adjusted R-Squared for the manual model excluding (including) the control variables is 2.56% (2.86%) as opposed to the automated model's 0.22% (0.39%). This means that the Adjusted R-Squared for the manual OLS models is 2.34% higher than the automated models without the control variables, and 2.47% higher when controls are added. The Vuong's (1989) test statistic provides strong evidence that, for regressing market returns on tone, manual models, as opposed to automated models, are closer to the true data generation process, both excluding (p=0.00) and including control variables (p=0.01). This finding rejects the null of H1 and suggests that a model with manual net tone score is better in explaining stock price movement than a model with automated net tone score.

For the fixed-effects regressions, the R-Squared for the manual model is 15.43% as opposed to the automated model's 13.36%, indicating that the R-Squared for the manual model is 2.07% higher than its automated counterpart. The Vuong (1989) test also suggests a rejection of the null of H1and indicates that the model with a manual net tone has greater power for explaining contemporaneous share price movements. These results are consistent with the findings of the OLS regressions.

In the OLS models including both manual and automated tone, both inclusive and exclusive of control variables, I find that the coefficient of TONE-M is positive and strongly significant, and very close in magnitude to that coefficient of TONE-M when modelled without TONE-A. TONE-A is negative but insignificant, unlike the magnitude and significance of its coefficient when modelled without TONE-M, suggesting that manual scoring is able to capture the tone for explaining share price movements, while automated scoring is not. The t-test for the difference between the coefficients of manual and automated tone supports this finding, both excluding (p=0.04) and including (p=0.04) control variables. In the FE models with individual manual and automated net tone scores, I find that the coefficients of both tones are positive and significant. When modelled together, I find that both tones are positive but only the manual tone is significant, i.e. the coefficient of TONE-M is 0.02 (p=0.00) while the coefficient of TONE-A is 0.01 (p=0.30). An estimate of the difference in coefficient reveals that the manual tone is larger but not significant (p=0.53).

[Insert Table 5]

It should be noted that in Table 5, I find that although the Adjusted R-Squared for OLS models and R-Squared for FE models are higher when manual net tone is used, the margin of increase in all cases range between 2-3%. The coefficients of TONE-M are marginally higher than those of TONE-A in comparable models. For instance, in the OLS models, excluding (including) the control variables, the coefficient of TONE-M is 0.03 (0.03) while the coefficient of TONE-A is 0.02 (0.02). A researcher has to determine whether this modest increase in the explanatory power and the tone coefficient magnitude for market returns is worth the additional time and labour needed when embarking on manual content analysis. As expected, I find that the size effect anomaly is not depicted in either the manual or automated models. In fact, the FE models reveal that firm size is positively associated with the share price reaction. In contrast, firms with greater risk have larger share price reaction in both OLS and FE estimations.

I subsequently re-estimate, using OLS regressions, eq. 1(a)-1(f) by adding a LOSS x TONE interaction term. This term interacts the 1/0 LOSS indicator variable with the tone variable(s) in the models. As a result, the coefficient of this interaction term represents the incremental market reaction to tone of loss firms, and the coefficient(s) of the original tone variable(s) represents the market reaction to the tone of profit firms. In un-tabulated results, I find, as expected, that this increases the magnitude of the tone coefficients for profit firms, while the interaction term is negative in all models. I continue to find, inclusive and exclusive of control variables, that the Adjusted R-Squared in manual models are modestly larger than in automated models. Vuong’s (1989) test results also support this, rejecting the null of H1, and suggesting that manual models have a greater explanatory power for market returns than automated models. The tests for difference in coefficients between manual and automated net tone scores for the combined manual and automated models are also qualitatively similar to the findings in Table 5.

The above finding should not necessarily be interpreted as manual content analysis being superior to automated content analysis for all linguistic features. Instead, the results in Table

5 suggest that the tone scored under manual content analysis has greater explanatory power for share price movement than the tone scored under automated content analysis. In this study the linguistic feature I examine is the tone of IMSs. It is possible that the superiority of manual content analysis method may not hold for a different linguistic feature, such as attributions of financial performance or readability.

4.2.2 Market Reaction to Tone Scores Decomposed into Positivity and Negativity

I now examine whether the market reactions to manual measures of positivity and negativity are different from their automated counterparts. This involves regressing CAR on comparable models of manual and automated measures of positivity and negativity, with and without control variables, and performing Vuong's (1989) test to see which model has greater explanatory power. An interesting observation from the descriptive statistics in Table 3 is the substantial difference in the dispersion of manual measures of positivity and negativity when compared to their automated counterparts. To elaborate, the standard deviation of manual net tone is only 1.8 times larger than that of automated net tone suggesting that any effect in the differences in skewness of TONE-A and TONE-M is unlikely to influence the overall results of Table 5. In contrast, the standard deviation of manual positivity (negativity) is 23 (48) times larger than automated positivity (negativity). Although manual and automated measures of positivity and negativity are computed using comparable formulas and allow the same possible range of values (0 to 1), the skewness of manual positivity and negativity are very different from their automated counterparts. There are at least two possible reasons for this difference in skewness. First, automated positivity and negativity scores are derived from the word frequencies in the entire IMS while manual positivity and negativity are derived from narratives scored selectively. Second, comparing the formulas of manual with automated positivity and negativity reveals that neutral narratives of financial performance in the denominator of manual positivity and negativity formulas, which were chosen by selective manual scoring, are compared to all words in an IMS document outside the Henry (2006) Lists of Positivity and Negativity that proxy for neutral words in the denominator of automated positivity and negativity formulas. To address this, I follow the approach of Henry and Leone (2009) and take the natural logarithms of 1 plus the manual and automated measures of positivity and negativity as a transformation method for the market reaction tests.

I focus primarily on the explanatory power of the models which determine how well the variations in CAR are explained by these measures. I begin by constructing positivity only and negativity only models of ordinary least square (OLS) regressions.

$$CAR = \alpha + \beta_1 POS\text{-}A + \varepsilon \quad \ldots (2a)$$

$$CAR = \alpha + \beta_1 NEG\text{-}A + \varepsilon \quad \ldots (2b)$$

$$CAR = \alpha + \beta_1 POS\text{-}M + \varepsilon \quad \ldots (2c)$$

$$CAR = \alpha + \beta_1 NEG\text{-}M + \varepsilon \quad \ldots (2d)$$

$$CAR = \alpha + \beta_1 POS\text{-}A + CONTROLS + \varepsilon \quad \ldots (2e)$$

$$CAR = \alpha + \beta_1 NEG\text{-}A + CONTROLS + \varepsilon \quad \ldots (2f)$$

$$CAR = \alpha + \beta_1 POS\text{-}M + CONTROLS + \varepsilon \quad \ldots (2g)$$

$$CAR = \alpha + \beta_1 NEG\text{-}M + CONTROLS + \varepsilon \quad \ldots (2h)$$

Then I create OLS models of manual and automated methods by putting the measures of positivity and negativity together, with and without control variables. This allows comparison between manual and automated models with both the components of the net tone—positivity and negativity—appearing in the same model, and therefore these models are simply tone decompositions of eq. 1(a), 1(b), 1(d) and 1(e) respectively.

$$CAR = \alpha + \beta_1 POS\text{-}A + \beta_2 NEG\text{-}A + \varepsilon \quad \ldots (2i)$$

$$CAR = \alpha + \beta_1 POS\text{-}M + \beta_2 NEG\text{-}M + \varepsilon \quad \ldots (2j)$$

$$CAR = \alpha + \beta_1 POS\text{-}A + \beta_2 NEG\text{-}A + CONTROLS + \varepsilon \quad \ldots (2k)$$

$$CAR = \alpha + \beta_1 POS\text{-}M + \beta_2 NEG\text{-}M + CONTROLS + \varepsilon \quad \ldots (2l)$$

Finally I combine manual and automated measures of positivity and negativity into single models:

$$CAR = \alpha + \beta_1 POS\text{-}A + \beta_2 NEG\text{-}A + \beta_3 POS\text{-}M + \beta_4 NEG\text{-}M + \varepsilon \quad \ldots (2m)$$

$$CAR = \alpha + \beta_1 POS\text{-}A + \beta_2 NEG\text{-}A + \beta_3 POS\text{-}M + \beta_4 NEG\text{-}M + CONTROLS + \varepsilon \quad \ldots (2n)$$

The control variables included in eq. 2 are the same variables in eq. 1(d) – 1(f) and are defined in Section 4.1. To supplement, I create fixed-effect regression models for manual and automated positivity and negativity measures, replacing 1/0 industry dummy variables with firm fixed-effects dummies:

$$CAR = \beta_1 POS\text{-}A + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2o)$$

$$CAR = \beta_1 NEG\text{-}A + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2p)$$

$$CAR = \beta_1 POS\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2q)$$

$$CAR = \beta_1 NEG\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2r)$$

$$CAR = \beta_1 POS\text{-}A + \beta_2 NEG\text{-}A + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2s)$$

$$CAR = \beta_1 POS\text{-}M + \beta_2 NEG\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2t)$$

$$CAR = \beta_1 POS\text{-}A + \beta_2 NEG\text{-}A + \beta_3 POS\text{-}M + \beta_4 NEG\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \dots (2u)$$

The results of eq. 2(a) – 2(u) are presented in Table 6. I first find that the Adjusted R-Squared of manual models are larger than comparable automated models in all cases, with and without the inclusion of control variables, and the results are generalizable across OLS and FE regressions. For OLS regressions, the increases in Adjusted R-Squared from automated negativity and positivity to comparable manual negativity and positivity is modest in nature, ranging between 1-2.5%; for FE regressions, the increases in R-Squared ranges between 1.5-2.5%. I also find that Vuong's (1989) test statistics provide strong evidence that the negativity and positivity tone scores computed under manual methods have greater explanatory power and are closer to the true data generation process in all cases, consistent with the direction of change in Adjusted R-Squared in OLS regressions and R-Squared in FE regressions. This rejects the null of H2a and H2b and indicates that models with manual measures of negativity as well as positivity yields greater power for explaining share price movements. Additionally, when automated positivity and negativity are compared to their manual counterparts in combined manual and automated models, I find, for OLS models, that manual positivity has larger reaction than automated positivity, both with and without controls, while the negativity comparisons are insignificant. In contrast, for FE models, I find that neither the coefficient of manual negativity or positivity is larger than their automated counterparts. Consistent with the findings of Table 5, I continue to find that firm size anomaly is not captured in these models. In particular, SIZE is positively associated with CAR in the FE estimations while depict no significant association in the OLS estimations.

The book-to-market ratio has a positive impact on market returns in both OLS and FE estimations.

[Insert Table 6]

I subsequently re-estimate the OLS models 2(a) – 2(n) in Table 6 by including two interaction terms for the LOSS indicator where I interact it with measures of positivity and negativity of each method. This now means that the coefficients of positivity and negativity measures are the market reactions of profit firms, while the interaction terms are the incremental market reaction for loss firms. The results, un-tabulated, indicate that all measures of positivity and negativity, in both manual and automated methods, yield a lower coefficient for loss firms. I find that loss firms also experience a greater reduction in share prices for negative words and narratives than profit firms. I continue to find that manual measures of negativity and positivity yield a marginally greater Adjusted R-Squared than comparable automated measures of negativity and positivity, consistent with the notion that manual measures of positivity and negativity are better in explaining contemporaneous share price movements.

4.3 Examining the Predictive Ability of Tone for Future Firm Performance

I now turn my attention to the second part of the study, which compares the predictive ability of manual and automated tones for future firm performance. For this, I regress a measure of *change* in firm performance on manual and automated net tone scores, with and without employing the control variables for regressing future firm performance on tone discussed in Section 4.1. Initially, I measure a firm’s future performance by CH_OP, which is the change in pre-exceptional operating profit between year t and t-1, deflated by total assets at the start of the year t. I devise the following ordinary least square (OLS) regression estimates:

$$CH\_OP = \alpha + \beta_1 TONE\text{-}A + \varepsilon \quad \dots (3a)$$

$$CH\_OP = \alpha + \beta_1 TONE\text{-}M + \varepsilon \quad \dots (3b)$$

$$CH\_OP = \alpha + \beta_1 TONE\text{-}A + \beta_2 TONE\text{-}M + \varepsilon \quad \dots (3c)$$

$$CH\_OP = \alpha + \beta_1 TONE\text{-}A + CONTROLS + \varepsilon \quad \dots (3d)$$

$$CH\_OP = \alpha + \beta_1 TONE\text{-}M + CONTROLS + \varepsilon \quad \dots (3e)$$

$$CH\_OP = \alpha + \beta_1 TONE\text{-}A + \beta_2 TONE\text{-}M + CONTROLS + \varepsilon \quad \ldots (3f)$$

The control variables included in eq. 3(d) – 3(f) are lagged operating profit change (LAGOP), earnings yield (EP), asset growth (AG), firm size (SIZE), book-to-market value of equity (BM), financial year buy-and-hold raw return (RET), eight 1/0 industry dummy variables (omitting 'INDUSTRIALS') and five 1/0 year dummy variables (omitting 'YEAR2008').

A potential criticism of the OLS estimates in eq. 3(d) – 3(f) is that they do not take into account the possibility of endogeneity bias. For example, it can be argued that firms with stronger 'disclosure incentives' or 'disclosure quality' and better 'monitoring' or 'accountability' tend to perform better. These firms may naturally report more positively in their IMSs, and they may also provide more details in their financial reports. I have not used a variable that directly measures 'disclosure quality' or 'monitoring' of firms in my OLS regression estimates. To deter the problems arising from potential omitted variable bias, I devise the following fixed-effects (FE) regressions to supplement eq. 3(d) – 3(f):

$$CH\_OP = \beta_1 TONE\text{-}A + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \ldots (3g)$$

$$CH\_OP = \beta_1 TONE\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \ldots (3h)$$

$$CH\_OP = \beta_1 TONE\text{-}A + \beta_2 TONE\text{-}M + FIRM\ FIXED\ EFFECTS + CONTROLS + \varepsilon \quad \ldots (3i)$$

The above FE regression models do not include 1/0 indicator variables for the industry.

The results of eq. 3(a) – 3(i) are presented in Table 7. For the OLS models, I find that, when manual models are compared to automated models, inclusive and exclusive of the control variables, manual models have marginally higher Adjusted R-Squared and Vuong's (1989) test statistic is highly significant suggesting that manual models are closer to the true data generating process. This rejects the null of H3 and suggests the tone computed under manual model is better in explaining future annual operating profit changes. When manual and automated net tone scores are put together in the same model, I find, with and without control variables, that the tone coefficient in the manual model is positive and highly significant, consistent with the findings of Davies et al. (2012); while the coefficient of tone in the automated model is negative and insignificant. The test of the difference of coefficients suggests that the manual net tone score is significantly better at capturing future annual changes in operating profit.

For the FE models, I find that the manual net tone score coefficient is positive and significant. This suggests that the manual tone can signal changes in future performance. On the other hand, I find that the automated net tone score coefficient is negative but insignificant. The Vuong (1989) test suggests that the manual tone has greater predictive ability for changes in future operating profit than its automated counterpart, rejecting the null of H3. When I include the manual and automated net tone scores in the same model, I continue to find that the manual tone score is positive and significant while the automated tone score is negative and insignificant at the 5% level. (p=0.01). Overall, I find that the results of the FE model is consistent with the findings of the OLS model and suggests that the manual tone has greater predictive ability for changes in future operating profits.

[Insert Table 7]

I additionally examine if my findings are generalizable to other measures of future firm performance. Therefore I select three additional measures of firm performance to replace CH_OP in eq. 3(a) – 3(i). Specifically, I use (a) CH_EPS, the percentage change in earnings per share in year t from year t-1, scaled by assets per share in year t-1 (b) CH_ROA, the percentage change in return on assets in year t from year t-1 and (c) CH_SALES, the percentage change in sales in year t from year t-1, scaled by total assets in year t-1. While CH_EPS and CH_ROA are derived from different measures of earnings, CH_SALES is derived from sales, is a firm fundamental regularly reported in an IMS (Schleicher and Walker, 2015). I re-estimate eq. 3(a) – 3(i) by alternatively replacing CH_OP with these three measures. All the explanatory variables in eq. 3(d) – 3(i) remain intact, with the exception of replacing LAGOP with LAGEPS, LAGROA and LAGSALES, which are *changes* in earnings per share, return on assets and sales in year t-1 from year t-2 respectively. I present the results of CH_EPS, CH_ROA and CH_SALES in Tables 8, 9 and 10 respectively.

For the OLS models, I find that in all cases that the manual method yields a marginally larger Adjusted R-Squared than the automated methods with and without the controls. For the OLS models of CH_ROA in Table 9, the Vuong’s (1989) test statistics suggest, with and without controls, that the tone under manual models are closer to the true data generating process of predicting future firm performance, rejecting the null of H3. For the OLS models of CH_EPS in Table 8, this is only true of the model without controls, while adding control variables renders the difference insignificant. For CH_SALES in Table 10, I find contrasting results on Vuong's (1989) statistics with and without controls: although the size and sign of the coefficients of TONE-A and TONE-M remain similar irrespective of whether the controls are added, I find that the tone under automated (manual) model is closer to the true data generating process for predicting future firm performance without (with) controls. In addition, the tests of difference in coefficients for OLS models in Tables 8 – 10 do not suggest that the tones are different in the majority of cases, especially when the control variables are added.

For the FE models, I find that CH_ROA in Table 9 and CH_SALES in Table 10 reject the null of H3 as the Vuong (1989) tests suggest that the manual tone scores have greater predictive power than their automated counterparts, while that for CH_SALES is negative but insignificant. However, I find that only the manual tone coefficient for the CH_ROA model is significantly positive. In contrast, the FE models in Table 8 suggest that the automated model has greater predictive power for changes in future firm performance. This rejects my non-directional null hypothesis of H3, but is not necessarily in direct contradiction with the Vuong (1989) tests of the OLS estimates (with control variables), which were inconclusive.

[Insert Tables 8, 9 & 10]

Overall, the findings in Tables 7 to 10 provide some indications that the tone of words and narratives can signal changes in future firm performance. The results are largely consistent across different measures of firm performance, with the exception of the estimates for changes in earnings per share. At the same time, I find that for the OLS estimates, the increases in Adjusted R-Squared from automated to manual tone is marginal, always ranging between 0.10%-1%, questioning whether the incremental benefits of manual scoring over automated scoring in terms of the tone's predictive power of future firm performance outweigh the costs, i.e. the time and labour necessary to manually read and score large volumes of financial disclosure. Likewise, for the FE models, I find that the difference in R-Squared between the manual and automated models is literally non-existent, ranging between 0.00%-0.01%. Given a virtual tie in model explanatory power, I suspect that the Vuong

(1989) tests for the FE models are largely driven by differences in the size and significance of the manual and automated net tone score coefficients.

5. Conclusion

In this study I compare the market reaction to the tone of the manual scoring with that of automated scoring and find that the explanatory power of the tone computed under manual method, for explaining market returns, is greater than the tone computed under automated returns. I further find that this greater explanatory power of manual over automated method holds when the net tone score is split into measures of negativity and positivity. I also find that the tone computed under manual content analysis has greater explanatory power for predicting future firm performance than the tone computed under automated content analysis. However, in all my estimations I observe that the increases in explanatory power of the tone from automated to manual method are modest to marginal in degree, and hence, the benefits of using manual content analysis for the tone to explain contemporaneous share price movements and to predict future performance must be weighed up with its costs, time consumption, and labour. For example, as a novice coder, it took me over five months to complete the manual scoring of all IMSs, including the time for pilot studies in the exploratory research phase. On the other hand, it took me only two days to learn how to handle WEBMATRIX and an additional single day to complete the automated scoring of all IMSs. An important question is whether the modest increase in explanatory power from automated to manual method is worth spending the additional time needed for manual scoring. Further, this result does not automatically indicate that manual content analysis is superior to automated content analysis for examining other linguistic features besides tone, such as attributions of performance or readability.

A related issue is perhaps to consider how many additional observations are needed in automated content analysis for its models to generate the same explanatory power as the manual models. I conduct an experimental analysis on this, using the OLS models of share price movements on tone, by making small subsamples for manual analysis and recording the resulting Adjusted R-Squared for various numbers of observations, and then directly comparing the Adjusted R-Squared with that of the full sample of automated analysis. I use four observation groups: samples of 50, 100, 250 and 500 IMSs. I repeat the regressions 5-10 times with different random samples for every observation group. I find that when 500 observations are used in manual content analysis, it occasionally yields a similar Adjusted R-

Squared as my automated content analysis of 1022 IMS observations. Given this to be true, I anticipate it would take well over two months to code 500 IMSs including the pilot studies, and so this finding reaffirms the argument that the additional time, labour and training required for manual content analysis potentially outweighs its incremental benefits.[11]

In my research design I explain that although manual content analysis only scores a subset of the text processed by WEBMATRIX, it is unlikely to compromise comparability of the two methods since all ignored statements are de-facto ‘neutral’ statements and neutral statements and words are not used to calculate the manual or automated net tone score. An alternative suggestion for research comparing manual and automated methods is to ensure the same corpus of texts is analyzed with both methods. Future capital market research can also use the evidence in my study to assert the relative benefits and demerits of manual as opposed to automated content analysis in accurately explaining share price movements as well as capturing other types of qualitative information in various types of financial disclosures such as preliminary earnings announcements, trading updates, or specific sections of an annual report.

[11] Two issues have to be considered for this experiment. First, the Vuong (1989) test cannot be used reliably in this experiment because it is an asymptotic test, and lacks the power for small to moderate sized samples, such as 50, 100 or even 200 observations (Clarke, 2007). Second, comparing Adjusted R-Squared of vastly different manual and automated method sample sizes in OLS models is not very reliable because Adjusted R-Squared not only depend on the number of observations and explanatory variables but also on the R-Squared value, which in turn depends on the explained and residual sum of squares, i.e. the specific values of the observations. As a result, when I use small subsamples of 50, 100 or 250 observations, the resulting Adjusted R-Squared are often spurious, and hence are likely unreliable for comparison.

Appendix 1 Coding Guideline for Manual Content Analysis

The appendix includes a step-by-step coding guideline for manual content analysis.

Manual Content Analysis Guideline and Coding Rules

1. Prepare an MS-Excel spreadsheet with the following column headings:
   - a. Name: name of the firm disclosing the IMS, as written in the IMS
   - b. Code: six-digit DataStream code
   - c. Year: year of IMS disclosure (2008, 2009, 2010, 2011, 2012 or 2013)
   - d. Event: 1 for first-quarter IMS and 2 for third-quarter IMS
   - e. Date: date of IMS disclosure, in format DD/MM/YYYY
   - f. POS: 1 if the tone of the statement is 'Positive', 0 otherwise.
   - g. NEU: 1 if the tone of the statement is 'Neutral', 0 otherwise.
   - h. NEG: 1 if the tone of the statement is 'Negative', 0 otherwise.
   - i. Sales Topic: 1 if the statement refers to a 'sales' topic, 0 otherwise.
   - j. Earnings Topic: 1 if the statement refers to an 'earnings' topic, 0 otherwise.
   - k. Cost Topic: 1 if the statement refers to a 'cost' topic, 0 otherwise.
   - l. Trading Topic: 1 if the statement refers to a 'trading' topic, 0 otherwise.
   - m. Results Topic: 1 if the statement refers to a 'results' topic, 0 otherwise.
   - n. GUS Topic: 1 if the statement refers to 'general unspecified statement' topic, 0 otherwise.
   - o. Quantitative: 1 if the statement uses number to indicate financial performance, 0 otherwise.

Each row contains one statement. A new row is used for every additional statement in the same IMS.

2. Read the IMS manually and underline with a pencil any statement on financial performance. Group statements have preference over segmental statements, segmental statements are ignored and not recorded in the presence of a group statement.
   - (i) Statement: A statement contains one particular piece of information on financial performance. It may encompass one complete sentence, multiple sentences, or part of a sentence.
   - (ii) Financial Performance Statement: A financial performance statement refers to six broad topics—'sales', 'earnings', 'cost', 'trading', 'results' and 'general unspecified statements' (GUS). Sales topics include code words such as 'sales', 'turnover', 'revenue', 'volume' and 'order book'. Earnings topics include code words such as 'earnings', 'profit', 'gross profit', 'net profit', 'EBIT', 'EBT', 'EPS', 'operating profit', 'income', and 'operating income'. Cost topics include code words such as 'cost', 'fixed cost', 'variable cost', 'tax', 'operating expenses' and 'administrative expenses'. Trading topics include code words such as 'trading' and 'business'. Results include code words such as 'results' or 'performance'. GUS topics include code words such as 'outlook', 'progress', 'success' and 'failure'. This list is used as a guide and not exhaustive; judgment is applied in classifying words into the six financial performance topics.
   - (iii) Group Statement: A group statement refers to the whole firm and not to any individual business, geographic or product segment. Code words for group statements include 'the group', 'the company', 'the firm', or name of the company. In the absence of specific segmental statements, all financial performance statements are considered as group statements. This list is a guide and not exhaustive.

3. Identify the tone of the statement. The tone can be either:
   - (i) POS: A positive statement is a clear or direct indication of improvement in financial performance;
   - (ii) NEU: A neutral statement represents (a) neither a distinctively positive nor distinctively negative message, (b) when performance is in line with expectations, (c) when status quo performance is preserved; or
   - (iii) NEG: A negative statement is a clear or direct deterioration in financial performance.

4. Complete columns (a) – (o) in the MS-Excel spreadsheet (as listed in Step 1) for the statement identified.

Appendix 2 Henry (2006) List of Positive and Negative Words in Automated Content Analysis

The appendix includes the list of positive and negative keywords from Henry (2006) used in automated scoring through WEBMATRIX in this study.

NEGATIVITY

disappoint disappoints disappointing disappointed disappointment risk risks risky threat threats threaten threatened threatening penalty penalties negative negatives negatively fail fails failed failing failure weak weakness weaknesses weaken weakens weakening weakened difficult difficulty hurdle hurdles obstacle obstacles slump slumps slumping slumped uncertain uncertainty uncertainties unsettled unfavorable downturn depressed down decrease decreases decreasing decreased decline declines declining declined fall falls falling fell fallen drop drops dropping dropped deteriorate deteriorates deteriorating deteriorated worsen worsens worsening worse worst low lower lowest less least smaller smallest shrink shrinks shrinking shrunk below under challenge challenges challenging challenged poor poorly

POSITIVITY

pleased delighted reward rewards rewarding rewarded opportunity opportunities enjoy enjoys enjoying enjoyed encouraged encouraging positive positives success successes successful successfully succeed succeeds succeeding succeeded accomplish accomplishes accomplishing accomplished accomplishment accomplishments strong strength strengths certain certainty definite solid excellent stellar good leading achieve achieves achieved achieving achievement achievements progress progressing deliver delivers delivered delivering leader leading up increase increases increasing increased rise rises rising rose risen double doubled doubles improve improves improving improved improvement improvements enhance enhances enhanced enhancing enhancement enhancements strengthen strengthens strengthening strengthened stronger strongest strongly better best more most above record high higher highest greater greatest larger largest grow grows growing grew grown growth expand expands expanding expanded expansion exceed exceeds exceeded exceeding beat beats beating

References

Abrahamson, E. and E. Amir. 1996. The information content of the president's letter to shareholders. *Journal of Business, Finance and Accounting* 23 (8): 1157-1182.

Baginski, S., J. Hassell and W. Hillison. 2000. Voluntary causal disclosures: Techniques and capital market reaction. *Review of Quantitative Finance and Accounting*15 (1): 371–389.

Baginski, S., J. Hassell and M. Kimbrough. 2004. Why do managers explain their earnings forecasts? *Journal of Accounting Research* 42 (1): 1–29.

Berelson, B. 1952. *Content Analysis in Communication Research*. New York: Free Press.

Clatworthy, M., and M. J. Jones. 2003. Financial reporting of good news and bad news: Evidence from accounting narratives. *Accounting and Business Research* 33 (3), 171-185.

Collins, D.W., S.P. Kothari, J. Shanken, and R.G. Sloan. 1994. Lack of timeliness and noise as explanations for the low contemporaneous return-earnings association. *Journal of Accounting and Economics* 18 (3): 289–324.

Davis, A.K., J.M. Piger, and L.M. Sedor. 2012. Beyond the numbers: Measuring the information content of earnings press release language. *Contemporary Accounting Research* 29(3): 845-868.

DeFond, M., M.Hung, and R. Trezevant. 2007. Investor protection and the information content of annual earnings announcements: international evidence. *Journal of Accounting and Economics* 43(1): 37-67.

Deloitte & Touche, 2007. *First IMpressionS: The First Year's Interim Management Statements*. London: The Creative Studio at Deloitte.

Doran, J.S., D.R. Peterson, and S.M. Price. 2010. Earnings conference call content and stock price: The case of REITs. *Working paper*.

Francis, J., D. Philbrick, and K. Schipper. 1994. Autumn. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32 (2): 137-164.

Francis, J., K. Schipper, and L. Vincent. 2002. Expanded disclosures and the increased usefulness of earnings announcements. *The Accounting Review* 77 (3): 515-546.

Gelb, D.S. and P. Zarowin. 2002. Corporate disclosure policy and the informativeness of stock prices. *Review of Accounting Studies* 7 (1): 33–52.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 91): 125-153.

Henry, E. 2006. Market reaction to verbal components of earnings press releases: Event study using a predictive algorithm. *Journal of Emerging Technologies in Accounting* 3 (1): 1-19.

Henry, E. 2008. Are investors influenced by how earnings press releases are written? *Journal of Business Communication* 45 (4): 363-407.

Henry, E. and A.J. Leone. 2009. *Unpublished work of Henry and Leone (2016) titled* 'Measuring qualitative information in capital markets research'.
Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1470807

Henry, E. and A.J. Leone. 2016. Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone. *The Accounting Review* 91 (1): 153-178.

Hutton, A. P., G. S. Miller, and D. J. Skinner. 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41 (5): 867-890.

Katz, S.B. 2001. Language and persuasion in biotechnology communication with the public: How not to say what you are not going to say and not say it. *AgBioFiorum* 4(2): 93-97.

Kothari, S.P., X. Li, and J. Short. 2009. The effect of disclosures by management, analysts, and financial press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review* 84:1639-1670.

Levin, I.P., S.L. Schneider, and G.J. Gaeth. 1998. All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes* 76 (2): 149-188.

Li. 2008. Annual report readability, current earnings and earnings persistence. *Journal of Accounting and Economics* 45 :221-247

Link, B. 2012. The struggle for a common interim frequency regime in Europe. *Accounting in Europe* 9 (2): 191-226.

Lopes, A.B. and M. Walker. 2012. Asset revaluations, future firm performance and firm-level corporate governance arrangements: new evidence from Brazil. *British Accounting Review* 44 (2): 53–67.

Loughran, T. and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance* 66 (1): 35-65.

Mahoney,W. and J. Lewis. 2004. *The IR Book*. Available on-line at http://www.ir-book.com.

Merkl-Davies, D.M. and N.M. Brennan. 2007. Discretionary disclosure strategies in corporate narratives: incremental information or impression management? *Journal of Accounting Literature* 27(1): 116-196.

Neuendorf, K. A. 2002. *The Content Analysis Guidebook.* Sage Publications.

Rutherford, B. 2005. Genre analysis of corporate annual report narratives: A corpus linguistics-based approach. *The Journal of Business Communication* 42 (4): 349-378.

Schleicher, T. And M. Walker. 2010. Bias in the tone of forward looking narratives. *Accounting and Business Research* 40 (4): 371-390.

Schleicher, T. and Walker, M. 2015. Are interim management statements redundant? *Accounting and Business Research* 45 (2): 229-255.

Simpson, S.D. 2014. Seven market anomalies investors should know. *Investopedia*.

Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance* 62 (3): 1139–1168.

Tetlock, P., M. Saar-Tsechansky, and S. Macskassy. 2008. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance* 63 (3): 1437-1467.

Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57 (2): 307–333.

Table 1 Sample Selection

The table illustrates the sample selection procedure. The sampling period spans 6 years, 2008-2013. 2008 is used as the year of sample determination, which had 668 firms in the FTSE All-Share Index as at 30 June. Eliminating: (a) all financial firms and (b) all non-financial firms publishing full quarterly results leave 324 non-financial firms disclosing an IMS for 2008. I randomly select 100 firms from this list, which can yield a maximum of 1200 IMSs during the 6 year period. I subsequently lose observations due to: (a) IMSs not available in the PI Navigator either because they were not disclosed by the firm in a particular period and / or missing from the PI Navigator and (b) collapse of a firm, resulting in a final 1022 IMSs in the sampling period. For manual content analysis, I score financial performance statements related to six broad topics: 'Sales', 'Earnings', 'Cost', 'Trading', 'Results' and General Unspecified Statements or 'GUS'. The number of positive, negative and neutral statements scored under each of these topics is given below. POS=positive statement, NEU=neutral statement, NEG=negative statement.

FIRM SAMPLE				
Firms in FTSE All-Share Index 30 June 2008				668
Less: Financial Firms				(305)
FTSE All-Share Index Non-Financial Firms 30 June 2008				363
Less: Non-Financial Firms releasing Quarterly Statements				(39)
FTSE All-Share Index Non-Financial Firms disclosing IMS				**324**
Randomly Selected Non-Financials from 30 June 2008				**100**
SIZE COMPOSITION IN SELECTED SAMPLE				
FTSE 100				15
FTSE 250				38
FTSE Small Cap				47
Total Firms				100
IMS SAMPLE				
Total Number of Firms				100
Maximum Possible IMS from Sample Firms				1200
Less: IMS lost due to collapse or delisting of firm				(69)
Less: IMS not disclosed by firms and / or missing from PI Navigator				(109)
Actual Number of IMS scored				**1022**
STATEMENT SAMPLE FOR MANUAL CONTENT ANALYSIS	POS	NEU	NEG	TOTAL
Statements on 'SALES'	964	92	405	1461
Statements on 'EARNINGS'	372	106	124	602
Statements on 'COST'	150	33	70	253
Statements on 'TRADING'	363	278	390	1031
Statements on 'RESULTS'	380	124	96	600
Statements on 'GENERAL UNSPECIFIED STATEMENTS (GUS)'	651	67	78	796
Total Statements	2880	700	1163	**4743**

Table 2 Manual Tone Scoring of Financial Performance Statements: Example

The table provides some examples of positive, neutral and negative narratives (performance statements) scored in manual content analysis. Words and / or phrases that played a key role in determining the tone (positive, neutral or negative) for manual content analysis are underlined.

Tone (Topic)	*Firm and IMS*	*Example*
Positive (GUS)	Lamprell plc (IMS published on 19.12.2012)	Moving in 2013, the Company will refocus on projects which are closer to the core business of Lamprell. In this way, the company can <u>leverage its historical strengths</u> and move forward <u>positively</u>.
Neutral (Results)	Future plc (IMS published on 29.7.2008)	While the market remains tough, Future is proving resilient and we will remain firmly <u>on track</u> for a satisfactory outturn for the full year.
Negative (GUS)	Greene King plc (IMS published on 28.1.2009)	The anticipated post-New Year <u>slowdown</u> has not, as yet, taken place, but we remain very cautious as to trading prospects for 2009, and we anticipate that the outlook for the rest of the year will remain very <u>challenging</u>.
Positive (Sales)	Ted Baker plc (IMS published on 19.11.2009)	Ted Baker, the British designer brand, is pleased to announce an 8.2% <u>increase</u> in Group revenue for the 13 week period from 16 August to 14 November (the 'period'), compared to the same period last year [...]
Neutral (Trading)	SSL International (IMS published on 28.1.2009)	Trading remains <u>in line</u> with expectations through the third quarter [...]
Negative (Earnings)	Centrica plc (IMS published on 11.5.2009)	Upstream profits are <u>adversely</u> impacted by lower commodity prices, <u>reducing</u> total Group operating profit

Table 3 Summary Statistics

The table presents summary statistics of the discrete and continuous variables used in the study from 1022 Interim Management Statements of 100 randomly selected non-financial FTSE All-Share Index firms during the period 2008-2013. CAR is cumulative abnormal return, which is the share price reaction over a three-day event (days t-1, t, t+1) centred on the announcement day of IMS obtained from the Perfect Information Navigator. For abnormal returns, I calculate daily market model adjusted returns, u_{it}, as $u_{it} = R_{it} - (\alpha_i + \beta_i R_{mt})$, where R_{it} is the return of firm *i* on day *t*, R_{mt} is the return of the FTSE All-Share Index on day *t* and where R_{it} and R_{mt} are calculated from DataStream Return Indices, RI. α_i and β_i are firm *i*'s estimated market model parameters calculated from the non-event period which runs from t-60 to t-10 and t+10 to t+60 relative to the IMS announcement day t=0. CAR is calculated as the sum of the daily market model adjusted returns, u_{it}, over the three-day event period (days t-1, t, t+1), such that $CAR_{it} = u_{it-1} + u_{it} + u_{it+1}$. LOSS is an indicator variable taking the value of 1 if pre-exceptional operating profit <0 at the start of the year t, and zero otherwise. SIZE is the natural log of market value of the equity of the firm at the beginning of the year t, calculated as: number of shares multiplied by share price. BM is the ratio of book-to-market value of the equity at the beginning of the year t. LENGTH is the number of words in an IMS document. RET is the financial year buy-and-hold raw returns. CH_ROA is change in return on assets in year t from year t-1, defined as operating profits divided by total assets in year t. CH_OP is change in pre-exceptional operating profits in year t from year t-1 deflated by total assets at the start of year t. CH_EPS is change in earnings per share in year t from year t-1 deflated by assets per share at the start of year t. EP is earnings yield, defined as operating profit divided by market value of equity at the start of year t. CH_SALES is annual change in sales in year t from year t-1, deflated by total assets at the start of year t. AG is asset growth, which is the percentage change in total assets at the start of year t. TONE-A is the net tone score from automated scoring, as (POSITIVE-A – NEGATIVE-A) / (POSITIVE-A + NEGATIVE-A) where POSITIVE-A and NEGATIVE-A are counted by WEBMATRIX as the number of words of Positivity and Negativity from the Henry (2006) List respectively. POS-A is the automated positivity score, calculated as number of positive words from Henry (2006) List scaled by the total number of words in the IMS document as counted by WEBMATRIX. NEG-A is the automated negativity score, calculated as number of negative words from Henry (2006) List scaled by the total number of words in the IMS document as counted by WEBMATRIX. TONE-M is the net tone score from manual scoring, calculated as (POSITIVE-M – NEGATIVE-M) / (POSITIVE-M + NEGATIVE-M) where POSITIVE-M and NEGATIVE-M are the total number of positive and negative statements scored in manual content analysis. POS-M is the manual positivity score, calculated as the number of positive statements divided by the sum of positive, neutral and negative statements scored. NEG-M is the manual negativity score, calculated as the number of negative statements divided by the sum of positive, neutral and negative statements scored. IMS OBS=1022.

Variable	*Mean*	*Std Dev*	*Median*	*Minimum*	*Maximum*	*Type*
CAR	0.0001	0.0856	0.0012	-0.5632	1.2239	Continuous
LOSS	0.1370	0.3437	0.0000	0.0000	1.0000	Discrete
SIZE	17.770	1.6695	17.563	10.437	22.592	Continuous
BM	0.5882	1.1881	0.4854	-12.500	25.000	Continuous
LENGTH	1008.0	819.00	771.00	107.00	9401.0	Continuous
CH_SALES	0.3017	7.4519	0.0306	-1.0000	337.35	Continuous
CH_EPS	0.5996	11.544	0.0156	-1.0000	561.84	Continuous
CH_ROA	0.4050	11.837	0.0000	-152.97	616.14	Continuous
CH_OP	1.4029	35.544	0.0104	-117.41	1698.0	Continuous
EP	1.1173	2.2675	0.9161	-10.740	65.530	Continuous
AG	0.2132	7.3963	-0.0073	-119.96	247.64	Continuous
RET	0.4636	7.2916	0.0191	-1.0000	351.07	Continuous
TONE-A	0.5893	0.2895	0.6364	-1.0000	1.0000	Continuous
POS-A	0.0280	0.0117	0.0275	0.0000	0.0721	Continuous
NEG-A	0.0067	0.0047	0.0059	0.0000	0.0404	Continuous
TONE-M	0.4552	0.5341	0.5000	-1.0000	1.0000	Continuous
POS-M	0.6157	0.2735	0.6667	0.0000	1.0000	Continuous
NEG-M	0.2240	0.2270	0.2000	0.0000	1.0000	Continuous

Table 4 Correlation Matrix

The table presents Spearman's rank correlations between the discrete and continuous variables used in the regression estimates of 1022 Interim Management Statements of 100 randomly selected non-financial FTSE All-Share Index firms during the period 2008-2013. Panel A presents the correlations between variables used to examine the association between share price reaction and IMS tone. Panel A also includes the variable LENGTH, for descriptive purpose. Panel B includes variables used to examine the association between IMS tone and future firm performance. In both panels, numbers are provided on the horizontal axis to identify the variables defined in the vertical axis.

CAR is cumulative abnormal return, which is the share price reaction over a three-day event (days t-1, t, t+1) centred on the announcement day of IMS obtained from the Perfect Information Navigator. For abnormal returns I calculate daily market model adjusted returns, u_{it}, as $u_{it} = R_{it} - (\alpha_i + \beta_i R_{mt})$, where R_{it} is the return of firm *i* on day *t*, R_{mt} is the return of the FTSE All-Share Index on day *t* and where R_{it} and R_{mt} are calculated from DataStream Return Indices, RI. α_i and β_i are firm *i*'s estimated market model parameters calculated from the non-event period which runs from t-60 to t-10 and t+10 to t+60 relative to the IMS announcement day t=0. CAR is calculated as the sum of the daily market model adjusted returns, u_{it}, over the three-day event period (days t-1, t, t+1), such that $CAR_{it} = u_{it-1} + u_{it} + u_{it+1}$.

LOSS is an indicator variable taking the value of 1 if pre-exceptional operating profit <0 at the start of the year t, and zero otherwise. SIZE is the natural log of market value of the equity of the firm at the beginning of the year t, calculated as: number of shares multiplied by share price. BM is the ratio of book-to-market value of the equity at the beginning of the year t. LENGTH is the number of words in IMS document.

RET is the financial year buy-and-hold raw return. CH_ROA is change in return on assets in year t from year t-1, defined as operating profits divided by total assets in year t. CH_OP is change in pre-exceptional operating profits in year t from year t-1 deflated by total assets at the start of year t. CH_EPS is change in earnings per share in year t from year t-1 deflated by assets per share at the start of year t. EP is earnings yield, defined as operating profit divided by market value of equity at the start of year t. CH_SALES is annual change in sales in year t from year t-1, deflated by total assets at the start of year t. AG is asset growth, which is the percentage change in total assets at the start of year t.

TONE-A is the net tone score from automated scoring, as (POSITIVE-A – NEGATIVE-A) / (POSITIVE-A + NEGATIVE-A) where POSITIVE-A and NEGATIVE-A are counted by WEBMATRIX as the number of words of Positivity and Negativity from the Henry (2006) List respectively. POS-A is the automated positivity score, calculated as number of positive words from Henry (2006) List scaled by the total number of words in the IMS document as counted by WEBMATRIX. NEG-A is the automated negativity score, calculated as number of negative words from Henry (2006) List scaled by the total number of words in the IMS document as counted by WEBMATRIX. TONE-M is the net tone score from manual scoring, calculated as (POSITIVE-M – NEGATIVE-M) / (POSITIVE-M + NEGATIVE-M) where POSITIVE-M and NEGATIVE-M are the total number of positive and negative statements scored in manual content analysis. POS-M is the manual positivity score, calculated as the number of positive statements divided by the sum of positive, neutral and negative statements scored. NEG-M is the manual negativity score, calculated as the number of negative statements divided by the sum of positive, neutral and negative statements scored. IMS OBS=1022. P-values are two-tailed.

PANEL A: IMS TONE AND SHARE PRICE REACTION											
	1	2	3	4	5	6	7	8	9	10	11
1. CAR	1.00	0.12	0.08	-0.12	0.19	0.19	-0.19	0.03	0.04	-0.01	-0.03
		0.00	0.01	0.00	0.00	0.00	0.00	0.36	0.19	0.69	0.31
2.TONE-A		1.00	0.51	-0.83	0.44	0.39	-0.42	-0.11	0.08	-0.08	-0.06
			0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.04
3. POS-A			1.00	-0.02	0.21	0.29	-0.15	-0.16	0.28	-0.06	0.12
				0.49	0.00	0.00	0.00	0.00	0.00	0.05	0.00
4. NEG-A				1.00	-0.40	-0.28	0.43	0.02	0.09	0.07	0.16
					0.00	0.00	0.00	0.46	0.00	0.02	0.00
5. TONE-M					1.00	0.80	-0.93	0.02	0.12	-0.10	0.02
						0.00	0.00	0.45	0.00	0.00	0.43
6. POS-M						1.00	-0.66	0.00	0.14	-0.07	0.10
							0.00	0.98	0.00	0.03	0.00
7.NEG-M							1.00	-0.06	-0.07	0.09	0.02
								0.05	0.03	0.01	0.52
8. LOSS								1.00	-0.21	0.10	0.17
									0.00	0.00	0.00
9. SIZE									1.00	-0.23	0.33
										0.00	0.00
10. BM										1.00	0.04
											0.23
11. LENGTH											1.00

PANEL B: IMS TONE AND FUTURE FIRM PERFORMANCE											
	1	2	3	4	5	6	7	8	9	10	11
1. CH_SALES	1.00	0.18	0.10	0.17	0.11	0.12	0.01	0.14	0.08	0.01	0.07
		0.00	0.00	0.00	0.00	0.00	0.46	0.00	0.01	0.86	0.00
2. CH_EPS		1.00	0.00	0.35	0.16	0.21	0.18	0.00	0.13	-0.09	0.25
			0.68	0.00	0.00	0.00	0.00	0.77	0.00	0.01	0.00
3. CH_ROA			1.00	-0.02	-0.06	0.02	-0.06	-0.13	-0.06	0.14	-0.01
				0.24	0.07	0.49	0.00	0.00	0.08	0.00	0.28
4. CH_OP				1.00	0.17	0.23	0.21	-0.01	0.07	-0.03	0.29
					0.00	0.00	0.00	0.28	0.03	0.39	0.00
5. TONE-A					1.00	0.44	-0.10	0.01	0.08	-0.08	0.19
						0.00	0.00	0.70	0.01	0.01	0.00
6. TONE-M						1.00	0.01	-0.03	0.12	-0.10	0.19
							0.86	0.37	0.00	0.00	0.00
7. EP							1.00	-0.01	-0.20	0.27	0.03
								0.62	0.00	0.00	0.03
8. AG								1.00	0.06	0.05	-0.03
									0.07	0.19	0.04
9. SIZE									1.00	-0.23	0.14
										0.00	0.00
10. BM										1.00	0.10
											0.00
11. RET											1.00

Table 5 Market Reaction to Tone Scores of Automated and Manual Methods

The table presents the ordinary least square (OLS) and fixed effects (FE) regressions of cumulative abnormal return, CAR on manual and automated net tone scores from 1022 Interim Management Statements of randomly selected 100 non-financial FTSE All-Share Index firms during the period 2008-2013. Panels A and C present the OLS regression estimates, without and with the control variables respectively. Panel E presents the FE regression estimates. 'M' refers to a model type where the variable of interest is derived by manual scoring. Panels B, D and F report the change in Adjusted R-Squared between manual only and automated only models and the test for the difference in the automated and manual net tone score coefficients in the combined models. Panel G includes the Vuong test of model preference for comparing manual and automated regressions. 'A' refers to a model type where the variable of interest is derived by automated scoring. CAR is cumulative abnormal return, which is the share price reaction over a three-day event (days t-1, t, t+1) centred on the announcement day of IMS obtained from the Perfect Information Navigator. For abnormal returns I calculate daily market model adjusted returns, u_{it}, as $u_{it} = R_{it} - (\alpha_i + \beta_i R_{mt})$, where R_{it} is the return of firm *i* on day *t*, R_{mt} is the return of the FTSE All-Share Index on day *t* and where R_{it} and R_{mt} are calculated from DataStream Return Indices, RI. α_i and β_i are firm *i*'s estimated market model parameters calculated from the non-event period which runs from t-60 to t-10 and t+10 to t+60 relative to the IMS announcement day t=0. CAR is calculated as the sum of the daily market model adjusted returns, u_{it}, over the three-day event period (days t-1, t, t+1), such that $CAR_{it} = u_{it-1} + u_{it} + u_{it+1}$. LOSS is an indicator variable taking the value of 1 if pre-exceptional operating profit <0 at the start of the year t, and zero otherwise. SIZE is the natural log of market value of the equity of the firm at the beginning of the year t, calculated as: number of shares multiplied by share price. BM is the ratio of book-to-market value of the equity at the beginning of the year t. TONE-A is the net tone score from automated scoring, as (POSITIVE-A – NEGATIVE-A) / (POSITIVE-A + NEGATIVE-A) where POSITIVE-A and NEGATIVE-A are counted by WEBMATRIX as the number of words of Positivity and Negativity from the Henry (2006) List respectively. TONE-M is the net tone score from manual scoring, calculated as (POSITIVE-M – NEGATIVE-M) / (POSITIVE-M + NEGATIVE-M) where POSITIVE-M and NEGATIVE-M are the total number of positive and negative statements scored in manual content analysis. YEAR2009, YEAR2010, YEAR2011, YEAR2012 and YEAR2013 are indicator variables to account for the year effect, taking the value of 1 if the IMS was disclosed in the relevant year and zero otherwise, and are relative to the 1/0 dummy variable YEAR2008. BASICMATERIALS, HEALTHCARE, CONSUMERGOODS, TECHNOLOGY, CONSUMERSERVICES, OILANDGAS, UTILITIES and TELECOMMUNICATIONS are respective indicator variables for the industry of the firm disclosing IMS, each denoting 1 if firm *i* is in the relevant industry and zero otherwise, and are all relative to the 1/0 dummy variable INDUSTRIALS (for OLS regressions only). Also included are fixed-effects dummy variables for 99 firms (for FE regressions only). All P-values are two-tailed.

	Dependent Variable: CAR					
	MODEL 1 (a)		MODEL 1 (b)		MODEL 1 (c)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL A: OLS - WITHOUT CONTROL VARIABLES						
INTERCEPT	-0.0090	0.1363	-0.0117	0.0008	-0.0091	0.1267
TONE-A	0.0164	0.0731			-0.0022	0.8265
TONE-M			0.0262	0.0000	0.0248	0.0000
Pr (F)	3.22	0.0731	27.79	0.0000	12.25	0.0000
OBS	1022		1022		1022	
R-Squared	0.0032		0.0266		0.0238	
Adj R-Squared	0.0022		0.0256		0.0218	

PANEL B: COMPARING ADJUSTED R-SQUARED AND T-TEST FOR COEFFICIENT DIFFERENCE IN OLS MODELS WITHOUT CONTROL VARIABLES				
1 (a) vs 1 (b)	Larger Adj. R-Squared	2.34%	1 (b)	M
1 (c)	Coeff Difference (p-value)	0.0394		M

(continued)	MODEL 1 (d)		MODEL 1 (e)		MODEL 1 (f)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL C: OLS - WITH CONTROL VARIABLES						
INTERCEPT	-0.0353	0.3003	-0.0301	0.3782	-0.0235	0.4886
TONE-A	0.0163	0.0878			-0.0022	0.8289
TONE-M			0.0267	0.0000	0.0255	0.0000
SIZE	0.0017	0.3623	0.0010	0.5921	0.0008	0.6728
BM	0.0072	0.0373	0.0090	0.0095	0.0086	0.0125
LOSS	0.0086	0.3382	0.0068	0.4513	0.0065	0.4678
BASICMATERIALS	-0.0026	0.8400	-0.0090	0.4867	-0.0090	0.4847
CONSUMERGOODS	0.0034	0.7163	0.0036	0.6986	0.0019	0.8370
CONSUMERSERVICES	0.0108	0.1145	0.0137	0.0470	0.0125	0.0681
OILANDGAS	0.0186	0.1982	0.0146	0.3168	0.0134	0.3503
UTILITIES	0.0010	0.9656	0.0037	0.8741	0.0024	0.9164
TECHNOLOGY	-0.0023	0.8357	0.0016	0.8835	0.0002	0.9885
TELECOMMUNICATIONS	0.0096	0.5038	0.0106	0.4560	0.0103	0.4720
HEALTHCARE	0.0036	0.8486	0.0062	0.7454	0.0050	0.7894
YEAR2009	-0.0108	0.2573	-0.0085	0.3732	-0.0063	0.5068
YEAR2010	-0.0152	0.1237	-0.0124	0.2092	-0.0121	0.2193
YEAR2011	-0.0187	0.0556	-0.0157	0.1102	-0.0153	0.1152
YEAR2012	-0.0202	0.0405	-0.0169	0.0875	-0.0164	0.0951
YEAR2013	-0.0216	0.1099	-0.0180	0.1755	-0.0182	0.1729
Pr (F)	1.23	0.2318	2.76	0.0002	2.38	0.001
OBS	1022		1022		1022	
R-Squared	0.0206		0.0448		0.0415	
Adj. R-Squared	0.0039		0.0286		0.0241	

PANEL D: COMPARING ADJUSTED R-SQUARED AND T-TEST FOR COEFFICIENT DIFFERENCE IN OLS MODELS WITH CONTROL VARIABLES						
1 (d) vs 1 (e)	Larger Adj. R-Squared			2.47%	1 (e)	M
1 (f)	Coeff. Difference (p-value)			0.0389		M

(continued)	MODEL 1 (g)		MODEL 1 (h)		MODEL 1 (i)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL E: FE - WITH CONTROL VARIABLES						
TONE-A	0.0304	0.0110			0.0132	0.3000
TONE-M			0.0288	0.0000	0.0233	0.0002
SIZE	0.0204	0.0019	0.0159	0.0145	0.0171	0.0098
BM	0.0213	0.0000	0.0242	0.0000	0.0225	0.0000
LOSS	0.0214	0.0492	0.0221	0.0436	0.0190	0.0799
FIRM FIXED EFFECTS	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	1.29	0.0294	1.54	0.0007	1.42	0.0051
OBS	1022		1022		1022	
R-Squared	0.1336		0.1543		0.1462	
PANEL F: COMPARING R-SQUARED AND TEST FOR DIFFERENCE IN TONE COEFFICIENTS IN FE MODELS WITH CONTROL VARIABLES						
1 (g) vs 1 (h)	Larger R-Squared			2.07%	1 (h)	M
1 (i)	Coeff. Difference (p-value)			0.5294		N/A
PANEL G: VUONG TEST OF MODEL PREFERENCE						
Models	*VUONG*	*Type*	*Z-Score*	*Pr>|z|*		
1 (a) vs 1 (b)	1 (b)	M	-5.2018	0.0000		
1 (d) vs 1 (e)	1 (e)	M	-2.8035	0.0051		
1 (g) vs 1 (h)	1 (h)	M	-4.7610	0.0000		

Table 6 Market Reaction to Automated and Manual Tone Scores Decomposed into Positive and Negative Measures

The table presents the ordinary least square (OLS) and fixed effects (FE) regressions of cumulative abnormal return, CAR on the positivity and negativity of manual and automated methods. It uses 1022 Interim Management Statements of 100 randomly selected non-financial FTSE All-Share Index firms during the period 2008-2013. 'M' refers to a model type where the variable of interest is derived by manual scoring. 'A' refers to a model type where the variable of interest is derived by automated scoring. CAR is cumulative abnormal return, which is the share price reaction over a three-day event (days t-1, t, t+1) centred on the announcement day of IMS obtained from the Perfect Information Navigator. For abnormal returns I calculate daily market model adjusted returns, u_{it}, as $u_{it} = R_{it} – (\alpha_i + \beta_i R_{mt})$, where R_{it} is the return of firm *i* on day *t*, R_{mt} is the return of the FTSE All-Share Index on day *t* and where R_{it} and R_{mt} are calculated from DataStream Return Indices, RI. α_i and β_i are firm *i*'s estimated market model parameters calculated from the non-event period which runs from t-60 to t-10 and t+10 to t+60 relative to the IMS announcement day t=0. CAR is calculated as the sum of the daily market model adjusted returns, u_{it}, over the three-day event period (days t-1, t, t+1), such that $CAR_{it} = u_{it-1} + u_{it} + u_{it+1}$. LOSS is an indicator variable taking the value of 1 if pre-exceptional operating profit <0 at the start of the year t, and zero otherwise. SIZE is the natural log of market value of the equity of the firm at the beginning of the year t, calculated as: number of shares multiplied by share price. BM is the ratio of book-to-market value of the equity at the beginning of the year t. POS-A is calculated as the natural logarithm of 1 plus the number of positive words from Henry (2006) List scaled by the total number of words in the IMS document as counted by WEBMATRIX. NEG-A is calculated as the natural logarithm of 1 plus the number of negative words from Henry (2006) List scaled by the total number of words in the IMS document as counted by WEBMATRIX. POS-M is the manual positivity score, calculated as the natural logarithm of 1 plus the number of positive statements divided by the sum of positive, neutral and negative statements scored. NEG-M is the manual negativity score, calculated as the natural logarithm of 1 plus the number of negative statements divided by the sum of positive, neutral and negative statements scored. YEAR2009, YEAR2010, YEAR2011, YEAR2012 and YEAR2013 are indicator variables to account for the year effect, taking the value of 1 if the IMS was disclosed in the relevant year and zero otherwise, and are relative to YEAR2008. BASICMATERIALS, HEALTHCARE, CONSUMERGOODS, CONSUMERSERVICES, OILANDGAS, UTILITIES, TELECOMMUNICATIONS, and TECHNOLOGY are respective indicator variables for the industry of the firm disclosing IMS, each denoting 1 if firm *i* is in the relevant industry and zero otherwise, and are all relative to INDUSTRIALS (for OLS regressions only). Also included are fixed-effects dummy variables for 99 firms (for FE regressions only). All P-values are two-tailed.

	Dependent Variable: CAR							
	MODEL 2 (a)		MODEL 2 (b)		MODEL 2 (c)		MODEL 2 (d)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL A: OLS - PARTIAL MODELS WIHTOUT CONTROL VARIABLES								
INTERCEPT	0.0172	0.1208	0.0199	0.0010	-0.0334	0.0000	0.0145	0.0002
POS-A	-0.0053	0.1260						
NEG-A			-0.0110	0.0004				
POS-M					0.0725	0.0000		
NEG-M							-0.0767	0.0000
Pr (F)	2.34	0.1260	12.56	0.0004	24.1	0.0000	26.09	0.0000
OBS	1022		1022		1022		1022	
R-Squared	0.0023		0.0123		0.0232		0.0250	
Adj. R-Squared	0.0013		0.0113		0.0222		0.0241	

PANEL B: COMPARING ADJUSTED R-SQUARED IN OLS PARTIAL MODELS WITHOUT CONTROL VARIABLES				
2 (a) vs 2 (c)	Larger Adj. R-Squared	1.00%	2 (c)	M
2 (b) vs 2 (d)	Larger Adj. R-Squared	2.28%	2 (d)	M

PANEL C: VUONG TEST OF MODEL PREFERENCE IN OLS PARTIAL MODELS WITHOUT CONTROL VARIABLES				
Models	*VUONG*	*Type*	*Z-Score*	*Pr>$\lvert z \rvert$*
2 (a) vs 2 (c)	2 (c)	M	-6.6265	0.0000
2 (b) vs 2 (d)	2 (d)	M	-3.8660	0.0001

(continued)	MODEL 2 (e)		MODEL 2 (f)		MODEL 2 (g)		MODEL 2 (h)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL D: OLS - PARTIAL MODELS WITH CONTROL VARIABLES								
INTERCEPT	-0.0420	0.2207	-0.0567	0.0981	-0.0494	0.1509	-0.0111	0.7466
POS-A	-0.0089	0.0265						
NEG-A			-0.0147	0.0000				
POS-M					0.0705	0.0000		
NEG-M							-0.0804	0.0000
SIZE	0.0042	0.0451	0.0047	0.0165	0.0011	0.5544	0.0014	0.4484
BM	0.0076	0.0270	0.0079	0.0213	0.0084	0.0155	0.0091	0.0088
LOSS	0.0102	0.2556	0.0134	0.1346	0.0084	0.3492	0.0052	0.5603
BASICMATERIALS	0.0018	0.8921	0.0086	0.5109	-0.0079	0.5416	-0.0093	0.4729
CONSUMERGOODS	0.0032	0.7365	0.0050	0.5897	0.0030	0.7489	0.0044	0.6416
CONSUMERSERVICES	0.0106	0.1217	0.0103	0.1297	0.0105	0.1279	0.0150	0.0301
OILANDGAS	0.0163	0.2598	0.0134	0.3501	0.0133	0.3612	0.0160	0.2703
UTILITIES	-0.0082	0.7228	-0.0032	0.8892	0.0048	0.8354	0.0003	0.9911
TECHNOLOGY	0.0001	0.9932	-0.0010	0.9249	-0.0003	0.9767	0.0014	0.9005
TELECOMMUNICATIONS	0.0113	0.4250	0.0116	0.4122	0.0115	0.4201	0.0117	0.4122
HEALTHCARE	0.0054	0.7745	0.0008	0.9658	0.0065	0.7340	0.0052	0.7833
YEAR2009	-0.0134	0.1563	-0.0076	0.4252	-0.0115	0.2284	-0.0075	0.4344
YEAR2010	-0.0151	0.1244	-0.0137	0.1614	-0.0148	0.1358	-0.0122	0.2176
YEAR2011	-0.0189	0.0535	-0.0171	0.0781	-0.0175	0.0749	-0.0154	0.1160
YEAR2012	-0.0209	0.0328	-0.0168	0.0875	-0.0192	0.0520	-0.0166	0.0928
YEAR2013	-0.0242	0.0706	-0.0226	0.0885	-0.0190	0.1537	-0.0175	0.1891
Pr (F)	1.35	0.1518	2.13	0.0048	2.43	0.0010	2.75	0.0002
OBS	1022		1022		1022		1022	
R-Squared	0.0226		0.0351		0.0396		0.0446	
Adj. R-Squared	0.0059		0.0186		0.0233		0.0283	

PANEL E: COMPARING ADJUSTED R-SQUARED IN OLS PARTIAL MODELS WITH CONTROL VARIABLES				
2 (e) vs 2 (g)	Larger Adj. R-Squared	1.74%	2 (g)	M
2 (f) vs 2 (h)	Larger Adj. R-Squared	2.24%	2 (h)	M

PANEL F: VUONG TEST OF MODEL PREFERENCE IN OLS PARTIAL MODELS WITH CONTROL VARIABLES

Models	*VUONG*	*Type*	*Z-Score*	*Pr>|z|*
2 (e) vs 2 (g)	2 (g)	M	-2.4304	0.0151
2 (f) vs 2 (h)	2 (h)	M	-3.3544	0.0008

(continued)	MODEL 2 (i)		MODEL 2 (j)		MODEL 2 (k)		MODEL 2 (l)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL G: OLS - COMPLETE MODELS WITHOUT AND WITH CONTROL VARIABLES								
INTERCEPT	0.0157	0.1539	-0.0098	0.4029	-0.0576	0.0940	-0.0250	0.4785
POS-A	0.0018	0.6545			-0.0016	0.7201		
NEG-A	-0.0119	0.0013			-0.0141	0.0003		
POS-M			0.0416	0.0277			0.0354	0.0672
NEG-M			-0.0502	0.0092			-0.0578	0.0036
SIZE					0.0050	0.0179	0.0011	0.5677
BM					0.0079	0.0206	0.0091	0.0091
LOSS					0.0136	0.1302	0.0063	0.4860
BASICMATERIALS					0.0091	0.4913	-0.0097	0.4502
CONSUMERGOODS					0.005	0.5923	0.0037	0.6894
CONSUMERSERVICES					0.0103	0.1308	0.0134	0.0539
OILANDGAS					0.0132	0.3589	0.0138	0.3443
UTILITIES					-0.0042	0.8557	0.0030	0.8967
TECHNOLOGY					-0.0007	0.9463	0.0009	0.9333
TELECOMMUNICATIONS					0.0117	0.4055	0.0119	0.4055
HEALTHCARE					0.0009	0.9611	0.0054	0.7762
YEAR2009					-0.0078	0.4106	-0.0077	0.4243
YEAR2010					-0.0137	0.1621	-0.0127	0.1996
YEAR2011					-0.0171	0.0785	-0.0155	0.1150
YEAR2012					-0.0168	0.0862	-0.0165	0.0943
YEAR2013					-0.0228	0.0861	-0.0171	0.1990
Pr (F)	6.38	0.0018	15.53	0.0000	2.02	0.0071	2.79	0.0000
OBS	1022		1022		1022		1022	
R-Squared	0.0124		0.0297		0.0352		0.0478	
Adj. R-Squared	0.0105		0.0277		0.0178		0.0306	

PANEL H: COMPARING ADJUSTED R-SQUARED IN OLS COMPLETE MODELS				
2 (i) vs 2 (j)	Larger Adj. R-Squared	1.72%	2 (j)	M
2 (k) vs 2 (l)	Larger Adj. R-Squared	1.28%	2 (l)	M

PANEL I: VUONG TEST OF MODEL PREFERENCE IN OLS COMPLETE MODELS				
Models	*VUONG*	*Type*	*Z-Score*	*Pr>|z|*
2 (i) vs 2 (j)-POSITIVITY	2 (j)	M	-6.3429	0.0000
2 (i) vs 2 (j)-NEGATIVITY	2 (j)	M	-3.8848	0.0001
2 (k) vs 2 (l)-POSITIVITY	N/A	M	-0.9055	0.3652
2 (k) vs 2 (l)-NEGATIVITY	2 (l)	M	-3.5091	0.0004

(continued)	MODEL 2 (m)		MODEL 2 (n)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL J: OLS - COMBINED MODELS WITHOUT AND WITH CONTROL				
INTERCEPT	0.0113	0.4193	-0.0502	0.165
POS-A	-0.0065	0.1462	-0.0087	0.0679
NEG-A	-0.0043	0.2927	-0.0064	0.1430
POS-M	0.0506	0.0107	0.0483	0.0168
NEG-M	-0.0348	0.0844	-0.0347	0.1016
SIZE			0.0040	0.0590
BM			0.0092	0.0071
LOSS			0.0116	0.1997
BASICMATERIALS			0.0016	0.9029
CONSUMERGOODS			0.0032	0.7285
CONSUMERSERVICES			0.0107	0.1184
OILANDGAS			0.0076	0.5990
UTILITIES			-0.0028	0.9028
TECHNOLOGY			0.0011	0.9192
TELECOMMUNICATIONS			0.0132	0.3479
HEALTHCARE			0.0016	0.9328
YEAR2009			-0.0042	0.6613
YEAR2010			-0.0116	0.2357
YEAR2011			-0.0141	0.1457
YEAR2012			-0.014	0.1535
YEAR2013			-0.0187	0.1581
Pr (F)	8.87	0.0000	2.86	0.0001
OBS	1022		1022	
R-Squared	0.0341		0.0546	
Adj. R-Squared	0.0302		0.0355	

PANEL K: TESTS OF DIFFERENCE IN POSITIVITY AND NEGATIVITY COEFFICIENTS IN OLS COMBINED MODELS			
Models	*T-Tests*	*P-value*	*Larger*
2 (m)	POS-A vs POS-M	0.0083	POS-M
2 (m)	NEG-A vs NEG-M	0.1592	N/A
2 (n)	POS-A vs POS-M	0.0100	POS-M
2 (n)	NEG-A vs NEG-M	0.2172	N/A

(continued)	MODEL 2 (o)		MODEL 2 (p)		MODEL 2 (q)		MODEL 2 (r)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL L: FE - PARTIAL MODELS WITH CONTROL VARIABLES								
POS-A	-0.0022	0.7331						
NEG-A			-0.0169	0.0005				
POS-M					0.0790	0.0000		
NEG-M							-0.0893	0.0000
SIZE	0.0251	0.0001	0.0214	0.0009	0.0185	0.0039	0.0151	0.0209
BM	0.0202	0.0000	0.0205	0.0000	0.0239	0.0000	0.0241	0.0000
LOSS	0.0221	0.0429	0.0220	0.0428	0.0248	0.0236	0.0196	0.0744
FIRM FIXED EFFECTS	YES	YES	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES	YES	YES
Pr (F)	1.23	0.0676	1.35	0.0131	1.49	0.0014	1.54	0.0007
OBS	1022		1022		1022		1022	
R-Squared	0.1274		0.1387		0.1505		0.1543	

			MODEL 2 (s)		MODEL 2 (t)		MODEL 2 (u)	
PANEL M: FE - COMPLETE MODELS WITH CONTROL VARIABLES								
POS-A			0.0038	0.5735			-0.0010	0.8840
NEG-A			-0.0176	0.0005			-0.0096	0.0857
POS-M					0.0395	0.0841	0.0350	0.1287
NEG-M					-0.0627	0.0081	-0.0432	0.0870
SIZE			0.0205	0.0018	0.0153	0.0190	0.0172	0.0095
BM			0.0208	0.0000	0.0244	0.0000	0.0222	0.0000
LOSS			0.0218	0.0451	0.0210	0.0564	0.0190	0.0820
FIRM FIXED EFFECTS			YES	YES	YES	YES	YES	YES
YEAR DUMMY			YES	YES	YES	YES	YES	YES
Pr (F)			1.34	0.0149	1.56	0.0005	1.44	0.0031
OBS			1022		1022		1022	
R-Squared			0.1390		0.1570		0.1510	

(continued)

PANEL N: COMPARING R-SQUARED IN FE MODELS				
2 (o) vs 2 (q)	Larger R-Squared	2.31%	2 (q)	M
2 (p) vs 2 (r)	Larger R-Squared	1.56%	2 (r)	M
2 (s) vs 2 (t)	Larger R-Squared	1.80%	2 (t)	M

PANEL O: VUONG TEST OF MODEL PREFERENCE IN FE MODELS

Models	*VUONG*	*Type*	*Z-Score*	*Pr>|z|*
2 (o) vs 2 (q)	2 (q)	M	-5.8489	0.0000
2 (p) vs 2 (r)	2 (r)	M	-3.8941	0.0000
2 (s) vs 2 (t)-POSITIVITY	2 (t)	M	-5.8654	0.0000
2 (s) vs 2 (t)-NEGATIVITY	2 (t)	M	-3.8726	0.0001

PANEL P: TESTS OF DIFFERENCE IN POSITIVITY AND NEGATIVITY COEFFICIENTS IN FE MODELS

Models	*T-Tests*	*P-value*	*Larger*
2 (t)	POS-A vs POS-M	0.1497	N/A
2 (t)	NEG-A vs NEG-M	0.2240	N/A

Table 7 The Predictive Ability of IMS Tone for Operating Profit

The table presents ordinary least square (OLS) and fixed effects (FE) regression coefficients and p-values of deflated change in pre-exceptional operating profit, CH_OP, on automated and manual IMS net tone scores, TONE-A and TONE-M on 1022 Interim Management Statements from a random sample of 100 firms during the period 2008-2013. CH_OP is the percentage change in operating profit in year t from year t-1, deflated by total assets at the end of the year t-1. LAGOP is CH_OP in the previous year. TONE-A is the net tone score from automated scoring in year t while TONE-M is the net tone score from manual scoring in year t. Earnings yield, EP, and book-to-market value of equity, BM, are defined as operating profit and start-of-year book value of equity, both deflated by market value of equity in year t-1. AG is asset growth, which is the percentage change in total assets at the start of the year t. SIZE is the natural log of market value of the equity of the firm at the start of the year, calculated as: number of shares multiplied by share price. RET is the financial year buy-and-hold raw returns. Also included are 1/0 indicator variables for 5 years omitting YEAR2008, for 8 industries relative to ‘INDUSTRIALS’ (for OLS regression only) and for 99 firms (for FE regression only). ‘M’ refers to manual scoring method, ‘A’ refers to automated scoring method. P-values are two tailed.

	Dependent Variable: CH_OP					
	MODEL 3 (a)		MODEL 3 (b)		MODEL 3 (c)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL A: OLS - WITHOUT CONTROL VARIABLES						
INTERCEPT	0.6696	0.0967	0.3205	0.1523	0.0672	0.0915
TONE-A	-0.0270	0.9642			-0.7134	0.2836
TONE-M			0.7306	0.0248	0.8938	0.0129
Pr (F)	0.00	0.9642	5.05	0.0248	3.11	0.0453
OBS	1022		1022		1022	
R-Squared	0.0000		0.0059		0.0073	
Adj. R-Squared	-0.0012		0.0047		0.005	
	MODEL 3 (d)		MODEL 3 (e)		MODEL 3 (f)	
PANEL B: OLS - WITH CONTROL VARIABLES						
INTERCEPT	7.7136	0.0010	7.3911	0.0014	7.5724	0.0012
TONE-A	0.3425	0.5969			-0.3903	0.5788
TONE-M			0.9531	0.0067	1.0374	0.0069
LAGOP	-0.0254	0.4576	-0.0268	0.4289	-0.0280	0.4110
EP	-0.0007	0.4250	-0.0009	0.3121	-0.0009	0.2950
AG	-0.0397	0.5306	-0.0370	0.5566	-0.0382	0.5451
SIZE	-0.3475	0.0053	-0.3538	0.0043	-0.3506	0.0049
BM	-0.2402	0.3151	-0.1968	0.4080	-0.1865	0.4374
RET	-0.6003	0.1020	-0.7057	0.0554	-0.7081	0.0556
INDUSTRY DUMMY	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	1.54	0.0627	1.93	0.0087	1.84	0.0123
OBS	1022		1022		1022	
R-Squared	0.0374		0.0464		0.0464	
Adj. R-Squared	0.0130		0.0223		0.0214	

(continued)	MODEL 3 (g)		MODEL 3 (h)		MODEL 3 (i)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL C: FE - WITH CONTROL VARIABLES						
TONE-A	-0.5595	0.4536			-1.4412	0.0757
TONE-M			0.8792	0.0151	1.0816	0.0059
LAGOP	-0.1230	0.0004	-0.1107	0.0015	-0.1239	0.0004
EP	-0.0019	0.3010	-0.0021	0.2424	-0.0023	0.2057
AG	-0.0245	0.7026	-0.0229	0.7221	-0.0188	0.7691
SIZE	-0.7320	0.0001	-0.7346	0.0001	-0.6986	0.0002
BM	0.4984	0.1306	0.4404	0.1820	0.5292	0.1078
RET	-1.1960	0.0034	-1.2275	0.0028	-1.2977	0.0015
FIRM FIXED EFFECTS	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	2.88	0.0000	2.79	0.0000	2.95	0.0000
OBS	1022		1022		1022	
R-Squared	0.2467		0.2399		0.2547	

PANEL D: TESTS OF COMPARISON				
3 (a) vs 3 (b)	Larger Adj. R-Squared	0.59%	3 (b)	M
3 (c)	Coeff. Difference (p-value)	0.0667		M
3 (d) vs 3 (e)	Larger Adj. R-Squared	0.93%	3 (e)	M
3 (f)	Coeff. Difference (p-value)	0.1223		N/A
3 (g) vs 3 (h)	Larger R-Squared	0.01%	3 (g)	A
3 (i)	Coeff. Difference (p-value)	0.0145		M

	VUONG	*Type*	*Z-Score*	*Pr>|z|*
3 (a) vs 3 (b)	3 (b)	M	-24.146	0.0000
3 (d) vs 3 (e)	3 (e)	M	-71.974	0.0000
3 (g) vs 3 (h)	3 (e)	M	-72.229	0.0000

Table 8 The Predictive Ability of IMS Tone for Earnings

The table presents ordinary least square (OLS) regression and fixed effects (FE) coefficients and p-values of deflated change in earnings per share, CH_EPS, on automated and manual IMS net tone scores, TONE-A and TONE-M on 1022 Interim Management Statements from a random sample of 100 firms during the period 2008-2013. CH_EPS is the percentage change in earnings per share in year t from year t-1, deflated by assets per share at the end of the year t-1. LAGEPS is CH_EPS in the previous year. TONE-A is the net tone score from automated scoring in year t. TONE-M is the net tone score from manual scoring in year t. Earnings yield, EP, and book-to-market value of equity, BM, are defined as operating profit and start-of-year book value of equity, both deflated by market value of equity in year t-1. AG is asset growth, which is the percentage change in total assets at the start of the year t. SIZE is the natural log of market value of the equity of the firm at the start of the year, calculated as: number of shares multiplied by share price. RET is the financial year buy-and-hold raw returns. Also included are 1/0 indicator variables for 5 years omitting YEAR2008, for 8 industries relative to 'INDUSTRIALS' (for OLS regression only) and for 99 firms (for FE regression only). 'M' refers to manual scoring method, 'A' refers to automated scoring method. P-values are two tailed.

	Dependent Variable: CH_EPS					
	MODEL 4 (a)		MODEL 4 (b)		MODEL 4 (c)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL A: OLS - WITHOUT CONTROL VARIABLES						
INTERCEPT	0.0266	0.8221	0.1111	0.0914	0.0331	0.7818
TONE-A	0.3239	0.0673			0.1530	0.4341
TONE-M			0.2507	0.0088	0.2212	0.0363
Pr (F)	3.36	0.0673	6.89	0.0088	3.82	0.0224
OBS	1022		1022		1022	
R-Squared	0.0039		0.0080		0.0090	
Adj. R-Squared	0.0028		0.0069		0.0066	
	MODEL 4 (d)		MODEL 4 (e)		MODEL 4 (f)	
PANEL B: OLS - WITH CONTROL VARIABLES						
INTERCEPT	0.1969	0.7662	0.2363	0.7189	0.1640	0.8051
TONE-A	0.3112	0.0916			0.1842	0.3597
TONE-M			0.2152	0.0325	0.1762	0.1095
LAGEPS	-0.0794	0.0357	-0.0867	0.0217	-0.0845	0.0260
EP	0.0000	0.9296	-0.0001	0.7596	-0.0001	0.8128
AG	-0.0017	0.9255	-0.0021	0.9084	-0.0015	0.9351
SIZE	-0.0172	0.6264	-0.0156	0.6581	-0.0170	0.6315
BM	-0.1318	0.0558	-0.1143	0.0968	-0.1209	0.0822
RET	0.5718	0.0000	0.5501	0.0000	0.5525	0.0000
INDUSTRY DUMMY	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	4.41	0.0000	4.51	0.0000	4.31	0.0000
OBS	1022		1022		1022	
R-Squared	0.1003		0.1022		0.1031	
Adj. R-Squared	0.0775		0.0796		0.0791	

(continued)	MODEL 4 (g)		MODEL 4 (h)		MODEL 4 (i)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL C: FE - WITH CONTROL VARIABLES						
TONE-A	0.4265	0.0441			0.3773	0.1027
TONE-M			0.1265	0.2161	0.0593	0.5957
LAGEPS	-0.3191	0.0000	-0.3219	0.0000	-0.3200	0.0000
EP	0.0005	0.3192	0.0005	0.3571	0.0005	0.3441
AG	-0.0099	0.5852	-0.0093	0.6080	-0.0096	0.5975
SIZE	-0.0807	0.1316	-0.0736	0.1696	-0.0793	0.1405
BM	-0.0330	0.7327	-0.0184	0.8484	-0.0303	0.7547
RET	0.3640	0.0019	0.3517	0.0028	0.3574	0.0025
FIRM FIXED EFFECTS	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	3.85	0.0000	3.81	0.0000	3.79	0.0000
OBS	1022		1022		1022	
R-Squared	0.3025		0.2995		0.3027	

PANEL D: TESTS OF COMPARISON				
4 (a) vs 4 (b)	Larger Adj. R-Squared	0.41%	4 (b)	M
4 (c)	Coeff. Difference (p-value)	0.7910		N/A
4 (d) vs 4 (e)	Larger Adj. R-Squared	0.21%	4 (e)	M
4 (f)	Coeff. Difference (p-value)	0.9760		N/A
4 (g) vs 4 (h)	Larger R-Squared	0.00%	4 (g)	A
4 (i)	Coeff. Difference (p-value)	0.2791		N/A

	VUONG	*Type*	*Z-Score*	*Pr>|z|*
4 (a) vs 4 (b)	4 (b)	M	-2.5889	0.0000
4 (d) vs 4 (e)	N/A	M	-1.3666	0.1718
4 (g) vs 4 (h)	4 (g)	A	1.9832	0.0473

Table 9 The Predictive Ability of IMS Tone for Return on Assets

The table presents ordinary least square (OLS) regression and fixed effects (FE) coefficients and p-values of deflated change in return on assets, CH_ROA, on automated and manual IMS net tone scores, TONE-A and TONE-M on 1022 Interim Management Statements from a random sample of 100 firms during the period 2008-2013. CH_ROA is the percentage change in return on assets in year t from year t-1, where return on assets is calculated as operating profits divided by total assets. LAGROA is CH_ROA in the previous year. TONE-A is the net tone score from automated scoring in year t. TONE-M is the net tone score from manual scoring in year t. Earnings yield, EP, and book-to-market value of equity, BM, are defined as operating profit and start-of-year book value of equity, both deflated by market value of equity in year t-1. AG is asset growth, which is the percentage change in total assets at the start of the year t. SIZE is the natural log of market value of the equity of the firm at the start of the year, calculated as: number of shares multiplied by share price. RET is the financial year buy-and-hold raw returns. Also included are 1/0 indicator variables for 5 years omitting YEAR2008, for 8 industries relative to 'INDUSTRIALS' (for OLS regression only) and for 99 firms (for FE regression only). 'M' refers to manual scoring method, 'A' refers to automated scoring method. P-values are two tailed.

	Dependent Variable: CH_ROA					
	MODEL 5 (a)		MODEL 5 (b)		MODEL 5 (c)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL A: OLS - WITHOUT CONTROL VARIABLES						
INTERCEPT	0.5675	0.1756	0.2349	0.3127	0.5843	0.1675
TONE-A	0.0274	0.9651			-0.6795	0.3260
TONE-M			0.7667	0.0235	0.9228	0.0135
Pr (F)	0.00	0.9651	5.15	0.0235	3.06	0.0473
OBS	1022		1022		1022	
R-Squared	0.0000		0.0060		0.0072	
Adj. R-Squared	-0.0012		0.0049		0.0049	
	MODEL 5 (d)		MODEL 5 (e)		MODEL 5 (f)	
PANEL B: OLS - WITH CONTROL VARIABLES						
INTERCEPT	8.2923	0.0007	7.9778	0.0009	8.1398	0.0008
TONE-A	0.4158	0.5367			-0.3479	0.6340
TONE-M			1.0070	0.0058	1.0821	0.0067
LAGROA	-0.0257	0.4695	-0.0272	0.4407	-0.0283	0.4241
EP	-0.0008	0.3878	-0.0010	0.2783	-0.0010	0.2649
AG	-0.0341	0.6046	-0.0313	0.6322	-0.0325	0.6206
SIZE	-0.3778	0.0035	-0.3839	0.0029	-0.3807	0.0033
BM	-0.2746	0.2690	-0.2280	0.3564	-0.2190	0.3801
RET	-0.6602	0.0836	-0.7701	0.0443	-0.7718	0.0447
INDUSTRY DUMMY	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	1.57	0.0537	1.97	0.0069	1.88	0.0100
OBS	1022		1022		1022	
R-Squared	0.0381		0.0473		0.0477	
Adj. R-Squared	0.0138		0.0233		0.0223	

(continued)	MODEL 5 (g)		MODEL 5 (h)		MODEL 5 (i)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL C: FE - WITH CONTROL VARIABLES						
TONE-A	-0.4579	0.5559			-1.3939	0.0988
TONE-M			0.9547	0.0112	1.1490	0.0050
LAGEPS	-0.1302	0.0004	-0.1180	0.0012	-0.1312	0.0003
EP	-0.0018	0.3280	-0.0021	0.2611	-0.0023	0.2246
AG	-0.0214	0.7487	-0.0190	0.7758	-0.0153	0.8179
SIZE	-0.7993	0.0000	-0.7988	0.0000	-0.7636	0.0001
BM	0.4309	0.2092	0.3773	0.2712	0.4634	0.1759
RET	-1.1982	0.0048	-1.2394	0.0037	-1.3061	0.0022
FIRM FIXED EFFECTS	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	2.80	0.0000	2.74	0.0000	2.88	0.0000
OBS	1022		1022		1022	
R-Squared	0.2446		0.2400		0.2530	

PANEL D: TESTS OF COMPARISON				
5 (a) vs 5 (b)	Larger Adj. R-Squared	0.61%	5 (b)	M
5 (c)	Coeff. Difference (p-value)	0.0786		M
5 (d) vs 5 (e)	Larger Adj. R-Squared	0.95%	5 (e)	M
5 (f)	Coeff. Difference (p-value)	0.1365		N/A
5 (g) vs 5 (h)	Larger R-Squared	0.00%	5 (g)	A
5 (i)	Coeff. Difference (p-value)	0.0179		M

	VUONG	*Type*	*Z-Score*	*Pr>|z|*
5 (a) vs 5 (b)	5 (b)	M	-30.485	0.0000
5 (d) vs 5 (e)	5 (e)	M	-77.293	0.0000
5 (g) vs 5 (h)	5 (h)	M	-89.964	0.0000

Table 10 The Predictive Ability of IMS Tone for Sales

The table presents ordinary least square (OLS) and fixed effects (FE) regression coefficients and p-values of deflated change in sales, CH_SALES, on automated and manual IMS net tone scores, TONE-A and TONE-M on 1022 Interim Management Statements from a random sample of 100 firms during the period 2008-2013. CH_SALES is the percentage change in sales in year t from year t-1, deflated by total assets at the end of the year t-1. LAGSALES is CH_SALES in the previous year. TONE-A is the net tone score from automated scoring in year t. TONE-M is the net tone score from manual scoring in year t. Earnings yield, EP, and book-to-market value of equity, BM, are defined as operating profit and start-of-year book value of equity, both deflated by market value of equity in year t-1. AG is asset growth, which is the percentage change in total assets at the start of the year t. SIZE is the natural log of market value of the equity of the firm at the start of the year, calculated as: number of shares multiplied by share price. RET is the financial year buy-and-hold raw returns. Also included are 1/0 indicator variables for 5 years omitting YEAR2008, for 8 industries relative to 'INDUSTRIALS' (for OLS regression only) and for 99 firms (for FE regression only). 'M' refers to manual scoring method, 'A' refers to automated scoring method. P-values are two tailed.

	Dependent Variable: CH_SALES					
	MODEL 6 (a)		MODEL 6 (b)		MODEL 6 (c)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL A: OLS - WITHOUT CONTROL VARIABLES						
INTERCEPT	0.0163	0.2191	0.0363	0.0000	0.0163	0.2239
TONE-A	0.0587	0.0031			0.0386	0.0784
TONE-M			0.0355	0.0010	0.0271	0.0217
Pr (F)	8.79	0.0031	10.97	0.001	7.03	0.0009
OBS	1022		1022		1022	
R-Squared	0.0102		0.0127		0.0161	
Adj. R-Squared	0.0091		0.0116		0.0141	
	MODEL 6 (d)		MODEL 6 (e)		MODEL 6 (f)	
PANEL B: OLS - WITH CONTROL VARIABLES						
INTERCEPT	0.0768	0.2822	0.0831	0.2414	0.0718	0.3161
TONE-A	0.0497	0.0126			0.0357	0.0987
TONE-M			0.0269	0.0132	0.0200	0.0906
LAGSALES	0.0623	0.0008	0.0638	0.0006	0.0613	0.0010
EP	0.0000	0.0747	0.0000	0.1395	0.0000	0.1075
AG	-0.0031	0.1136	-0.0032	0.1028	-0.0031	0.1176
SIZE	-0.0038	0.3131	-0.0032	0.4042	-0.0038	0.3243
BM	-0.0147	0.0448	-0.0132	0.0723	-0.0139	0.0596
RET	0.0262	0.0201	0.0243	0.0326	0.0245	0.0310
INDUSTRY DUMMY	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	5.43	0.000	5.39	0.0000	5.31	0.0000
OBS	1022		1022		1022	
R-Squared	0.1208		0.1197		0.1241	
Adj. R-Squared	0.0985		0.0975		0.1007	

(continued)	MODEL 6 (g)		MODEL 6 (h)		MODEL 6 (i)	
Variables	*Coeff*	*P-value*	*Coeff*	*P-value*	*Coeff*	*P-value*
PANEL C: FE - WITH CONTROL VARIABLES						
TONE-A	-0.1010	0.5652			0.0007	0.9972
TONE-M			-0.1263	0.1354	-0.1286	0.1650
LAGEPS	0.0350	0.8056	0.0386	0.7850	0.0388	0.7856
EP	0.0014	0.0010	0.0014	0.0007	0.0014	0.0007
AG	0.0137	0.3622	0.0129	0.3893	0.0131	0.3871
SIZE	0.0112	0.8004	0.0060	0.8914	0.0064	0.8847
BM	0.0902	0.2436	0.0862	0.2626	0.0875	0.2593
RET	-0.1888	0.0484	-0.1773	0.0638	-0.1785	0.0636
FIRM FIXED EFFECTS	YES	YES	YES	YES	YES	YES
YEAR DUMMY	YES	YES	YES	YES	YES	YES
Pr (F)	2.08	0.0000	2.12	0.0000	2.08	0.0000
OBS	1022		1022		1022	
R-Squared	0.1963		0.1985		0.1986	

PANEL D: TESTS OF COMPARISON				
6 (a) vs 6 (b)	Larger Adj. R-Squared	0.25%	6 (b)	M
6 (c)	Coeff. Difference (p-value)	0.6918		N/A
6 (d) vs 6 (e)	Larger Adj. R-Squared	0.10%	6 (d)	A
6 (f)	Coeff. Difference (p-value)	0.5796		N/A
6 (g) vs 6 (h)	Larger R-Squared	0.00%	6 (h)	M
6 (i)	Coeff. Difference (p-value)	0.5945		N/A

	VUONG	*Type*	*Z-Score*	*Pr>|z|*
6 (a) vs 6 (b)	6 (a)	A	6.0699	0.0000
6 (d) vs 6 (e)	6 (e)	M	-3.1957	0.0014
6 (g) vs 6 (h)	6 (e)	M	-3.5266	0.0004