Publication information

Title	Analyst Report Readability
Author(s)	Hsieh, Chia-Chun; Hui, Kai Wai
Source	American Accounting Association Annual Meeting, Denver, Colorado, USA , 6-10 August 2011
Version	Pre-published version
DOI	Nil
Publisher	Nil

Copyright information

© 2011 The Author(s)

Notice

This version is available at HKUST Institutional Repository via

<http://hdl.handle.net/1783.1/51253>

If it is the author's pre-published version, changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published version.

Analyst Report Readability

Chia Chun Hsieh
Department of Accounting
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong
achsieh@ust.hk

Kai Wai Hui^{*}
Department of Accounting
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong
ackw@ust.hk

August 2011

* We acknowledge the financial support from Hong Kong Research Grants Council (project No. 641310). We thank Philip Berger for his insightful comments. We also benefited from Kin Lo, Joy Begley, Sandra Chamberlain, Gilles Hilary, Peter Clarkson, and workshop participants at CAAA 2011 Conference, University of British Columbia and INSEAD. All errors are our own.

Analyst Report Readability

Abstract

This paper studies three research questions regarding analyst forecast reports' readability: (1) whether readability of analyst reports affects investment decisions; (2) whether the way analysts explain the reasoning of their research (i.e. how they tell their story) relates to accuracy of earnings forecasts; and (3) whether readability of analyst reports relates to the underlying characteristics of the analysts. We find that the market reacts more favorably to reports that revise earnings forecasts upward and are easier to read. Consistent with this finding, upward revisions have smaller forecast errors in relatively more readable reports. We also find that security analysts with longer experience, who follow more firms, tend to issue more readable reports. Upward revisions issued by analysts from smaller brokerage houses and analysts having better records of forecast accuracy are more readable. Reports are also more readable when they are issued immediately after firms' public disclosures or management forecasts. Overall, readability complements the information of reports containing upward revisions, and indicates better forecast accuracy.

1. Introduction

Analyst forecast reports contain detailed information regarding analysts' views of future profitability of companies. Asquith, Mikhail and Au (2005) suggested that disclosure of research logic and reasoning behind an analyst report is critical for investors to better understand firm value. One important feature of analyst reports that reflects value relevant information, as well as incentives of authors, is readability. For example, Li (2008) provided empirical evidence that managers use less complicated language in annual reports when presenting more persistent positive earnings. This paper examines whether linguistic features of analyst reports contain value relevant information. Specifically, we study three research questions: (1) whether the readability of analyst reports affects investment decisions; (2) whether the way in which analysts explain the reasoning of their research (i.e. how they tell their story) relates to the accuracy of earnings forecasts; and (3) whether readability of analyst reports relates to the underlying characteristics of individual analysts. That is, what are the determinants of analyst reports' readability?

We argue that the level of readability positively reflects an analyst's willingness to explain their forecasts, and that analysts issuing accurate forecasts are more likely to provide readable reports. Security analysts are important information intermediaries who facilitate communication between investors and managements (Healy and Palepu, 2001). Prior studies suggest that analysts' career outcomes are linked to the quality of their forecasting work, assessed by ex-post verification of forecast accuracies (Mikhail, Walther and Willis, 1999; Clement and Tse, 2003; Hong and Kubik, 2003). Analyst reports contain vital information and the reasoning behind forecast revisions. A well-written report, therefore, demonstrates how the forecast is formulated, and the analyst's ability to gather and interpret information. On the other hand, a poor forecasting work is also visible to the market if the reasoning behind an inaccurate forecast is published in a more readable report. Therefore, only analysts

with quality information to provide accurate forecasts are willing to provide readable reports. Readable reports better disseminate analysts' work to investors and, therefore, enhance analysts' reputation (for issuing accurate forecasts) but they also increase penalties for any inaccurate forecasts. As a result, we hypothesize that readability influences investors' reactions to analyst reports as it indicates more accurate analyst forecasts. In addition, since analysts who make more accurate forecasts are more willing to improve the readability of their reports, we hypothesize a positive association between readability and accuracy, as well as a positive relationship between readability and analyst reputation (for accurate forecasts).¹

To shed light on how analysts' incentives influence report readability, we further study the different relationships between report readability and informativeness of reports carrying good news (upward revision) and bad news (downward revision). Previous literature suggests that analysts have different incentives for issuing upward and downward revisions. For example, Ke and Yu (2006) suggested that analysts issue optimistic forecasts at the beginning of the year to obtain (more) information from managements and then "walk down" the forecasts so that the management can beat consensus forecasts of earnings. O'Brien, McNichols and Lin (2005) provided empirical evidence that affiliated analysts are more likely to issue good news as they are generally not willing to issue bad news forecasts. We, therefore, expect that consistent with managers' incentive to increase the credibility of good news forecasts (Hutton, Miller and Skinner, 2003), analysts are also more likely to improve credibility of upward revision forecasts, so that the good news has a strong influence on investors beliefs. We expect that if readability is a credible signal of accuracy, then readability has greater influence on investment decisions when analyst reports contain upward revisions rather than downward revisions. We further hypothesize that forecast accuracy is more related to readability when the forecasts result in upward revision of future

¹ Prior studies suggest that markets have stronger response towards analyst forecast revisions when the forecast is expected to be more accurate (Abarbanell et al., 1995; Stickel, 1992; Park and Stice, 2000; Gleason and Lee, 2003; and Clement and Tse, 2003).

earnings expectations. In addition, analyst reputation for accurate forecasts also has a strong association with readability when analysts issue upward revision forecasts.

To measure readability, we follow Li (2008) and other previous studies, using linguistic measures to define the complexity and ambiguity of a report (e.g., Fog index²). We first study the influence of readability on investment decisions in terms of stock returns around the time of issuance of a report. Second, we examine the association between readability of analyst reports and the level of (ex-post forecast) accuracy. Third, we study the association between readability and characteristics of the analysts and the quality of financial disclosures (i.e. public information) by firms under analysis.

In the first and second parts of the study, we test whether market reaction and ex-post forecast accuracy vary across different levels of readability. Using a sample of 2,930 analyst reports from the high-tech industry, in terms of three-day abnormal return around the time of report issuance, we find that coefficient of interaction between forecast revision and readability is significantly greater when the revision is upward, and when the report is more readable. We further explore the cross-section variation of the results. Specifically, we find the result is stronger when the forecast horizon is short and expected forecast accuracy is high, this is consistent with the notion that analysts issue readable reports when they have better confidence of forecast accuracy. We also find the stronger results when investors' demand of information and the price impact of forecast is greater, which is consistent with analyst having stronger reputation concern when their forecast is more important to the investors.

We further show that accuracy is significantly associated with report readability when reports revise forecasts upward. This is consistent with the expectation that readability complements the accuracy of forecasts revisions, and analysts are more willing to make their reports readable to reinforce accurate upward revision forecasts. On the other hand, we don't

² We also use two other commonly used measures in linguistics, Flesch and Flesch-Kincaid, to supplement the Fog index.

find similar results when analyst reports make downward revisions. This is consistent with prior evidence to the effect that investors, especially large traders, have better understanding of analysts' disincentives to lower earnings forecasts, are less skeptical about bad news reports and are more willing to trade on downgrades and sells (Mikhail, Walther and Willis 2007), which mitigates the need to make reports more readable to convey credibility.

In the third part of the study, our empirical findings indicate that analysts with broader coverage, high forecast frequency and longer firm specific experiences issue more readable reports. This finding is consistent with the notion of readability being a credible indication of analyst ability to issue accurate forecasts. We find that reports issued immediately after earnings announcements and management forecasts are more readable, which is consistent with the notion that readability is associated with better quality information input in forecasting (Lehavy, Li and Merkley, 2011). Similar to Li (2008), we also find that factors such as firm growth, mergers and acquisitions, number of segments and firm age influence readability.

To examine the different incentives for improving the readability of upward and downward revision reports, we further segregate our sample on the basis of different types of revisions of reports. Consistent with our expectation that analyst ability to make accurate forecasts is more likely to be associated with readability in upward revision reports, we find that reports of analysts with good historical records of issuing accurate forecasts are more likely to be more readable when the revisions are upward. When examining other determinants of readability, we find that smaller brokerage houses are more likely to issue readable reports only in case of upward revisions and not in case of downward revision reports, and that readability of only downward revision reports is influenced by earnings volatility and firm age.

Overall, our empirical evidence is consistent with the finding that readability matters more when reports convey upward revisions to the market. Readability of forecast reports is driven by analysts' ability to forecast accurately and the quality of information provided by managements. Readable reports have better forecast accuracy and market participants assign more weight to such reports when responding they carry good –news.

This study contributes to financial analysis literature by examining analyst reports, a product publicly available to investors but poorly understood by researchers. To the best of our knowledge, our study is among the first ones to examine whether readability of analyst reports affects the information content that goes beyond earnings forecasts or stock recommendations. This also contributes to literature on the value of analysts as information intermediaries. As readability directly affects an investor's understanding of the information provided in a report, it is indicative of whether the analyst has exerted effort to “promote or sell” the forecast and the recommendation of the product through a more readable document. By studying this linguistic feature of analyst reports, we find that the quality of writing associates positively with analysts' research effort and indicates better forecast accuracy.

The second contribution is analysis of whether report readability implies some important information that the market recognizes and reacts to. Readability reflects the difficulty (or ease) of understanding the information, which directly affects investors' interpretation of forecasts. By showing that readability of analyst reports affects investors' perceptions of forecasts, this study extends the existing research on market reaction to analyst reports. It provides empirical evidence that the quality of writing of analyst reports influences interpretation of information contained in the report by investors and, therefore, provides a better understanding of behavior of market participants.

In addition to examining these two issues, this paper also makes a contribution by incorporating linguistic research on text readability into analyst research. Asquith et al. (2005)

suggested that analyst reports provide useful information to investors whereas Li (2008) provided empirical evidence of the importance of readability of annual reports, and posited that studying the readability of analyst reports allows us to understand analyst reports from a different angle. Li et al. (2011) further extend the studies on financial reports' readability by studying the association between readability of 10K filings and analyst behaviors. As previous work in this area has not considered measuring the role of linguistics in financial analysis research, we know little about whether linguistic factors affect the role of analysts in financial markets. We explore the determinants of the language that analysts use in their reports and whether these linguistic features are associated with the underlying characteristics of the analysts, their reports and the firms. Our findings contribute to linguistic literature, as well as literature on financial analysts.

The remainder of this paper is organized as follows. Section 2 discusses prior literature and develops hypotheses. Section 3 describes research plan and methodology. Empirical findings are discussed in Section 4, and Section 5 concludes.

2. Literature Review and Hypothesis Development

Financial analysts have been the subject of much attention in literature on finance and accounting. Topics that have been studied comprehensively include forecast accuracy and analyst characteristics (Clement, 1999; Clement and Tse, 2003), analyst forecasts walking-down through the year (Ke and Yu, 2006), and the impact of analyst affiliation on forecasting performance (Barber, Lehavy, McNichols and Trueman, 2003; 2007; Cliff, 2007). Although these studies have generated abundant findings, which have allowed better understanding of the role of analysts in the market, their focus has been primarily on two final outputs of analysts: earnings forecasts and stock purchase recommendations. Several relatively more recent studies have examined the informativeness of analyst reports. For example, Kothari, Li

and Short et al. (2009) examined the tone (favorable versus unfavorable opinion) of analyst reports and Asquith et al. (2005) tested their specific information components (contents). In contrast to these studies, which focus on *what* information is communicated, this paper studies analyst reports from a different angle, using different tools, and focuses on *how* the information is communicated (i.e. the readability of the report).³ The study of analyst reports readability represents the hitherto unanswered question: whether the quality of writing of the report reflects forecast accuracy and whether it helps investors understand the forecasts/recommendations in the report.

In the first part of empirical tests, we study market reaction to reports. We expect report readability to influence market reaction because report readability is a positive indication of accuracy. In other words, only analysts having quality information to predict future earnings accurately issue readable reports. Prior studies suggest that investors value characteristics that indicate the accuracy of analyst forecasts (Clement and Tse, 2003). Analysts' career outcomes are also linked with ex-post verifiability of accuracy of their reports. For example, Mikhail et al. (1999) provided empirical evidence that more experienced analysts provide forecasts with better accuracy. Hong and Kubik (2003) suggested that analysts with better records of forecast accuracy are more likely to move up to a higher status brokerage house. It is, therefore, crucial for security analysts to establish their reputation for accurate forecasts.

Analyst reports include the reasoning behind the conclusion(s) and, therefore, provide vital information on whether the ex-post verified accuracy of a forecast can be ascribed to the ability of the analyst. By definition, a more readable report means the reader can better understand and process the information. This implies that investors have better understanding (lower information processing cost) of the logic behind analyst forecast or revision of earlier

³ Note that while analysts can be positive or negative in tone, this does not imply that both types have the same readability.

forecasts or expectations (e.g., Bloomfield, 2002; Lehavy et al., 2011). A well-written report, therefore, results in a more precise assessment of analyst ability to forecast accurately. For example, if the report is well-written, investors can understand that the EPS revision is because of, say, a favorable industry outlook. When a forecast is ex-post found to be accurate, investors have perceive the analyst's forecasting ability to be high. On the other hand, if the forecast is inaccurate, a well-written report reveals an analyst's inability to collect proper information and investors, therefore, discount that analyst's forecasting ability.

As a result, more (less) readable reports lead to greater reward (penalty) if the forecast is ex-post found accurate (inaccurate). To avoid the additional penalty of inaccurate forecasts and maximize the benefit of accurate forecasts, only analysts providing accurate forecasts are willing to provide readable reports. Readable reports, therefore, are associated with better ex-post forecast accuracy. Market participants, as a result, assign greater weight to forecast revisions (i.e., a bigger response coefficient), when updating their valuation based on the forecast revision (Clement and Tse, 2003). Overall, we predict that report readability complements the effect of information contained in analyst report and signals better accuracy, therefore, there is a positive association between report readability and stock return sensitivity to forecast revisions.⁴

However, we expect the relationships between returns and report readability of upward and downward revision forecasts to be different, because of a stronger concern of analysts' reputation to issue accurate forecasts in an optimistic report. First, analysts' incentive to improve report-writing may differ according to the positive or negative signals intended to be communicated by the report. Francis and Philbrick (1993) argued that analysts

⁴ The above discussion is based on the implicit assumption that writing of the report complements the underlying information of the report. However, it might be the case that readability has a substitution relationship with forecast information. For example, more capable analysts are good at providing precise forecasts, so a detailed report may be less necessary. In which case, an analyst with precise information would not try to issue a readable report, given that the value added to the information is already contained in earnings forecasts and stock recommendations.

issue optimistic earnings forecasts to maintain good relations with managements. O'Brien et al. (2005) provided empirical evidence that affiliated analysts are more likely to issue reports carrying good-news. Hong and Kubik (2003) suggested that after controlling for forecast accuracy, forecast optimism also leads to favorable job separation. Mikhail, Walther and Willis (2007) provided empirical evidence that compared with smaller traders, sophisticated market participants react less to favorable recommendation revisions but react strongly to unfavorable revisions. As a result, if analysts issue upward revision forecasts to create positive market reactions that favor the management, analysts have a greater incentive to improve the readability of their reports than when making downward revisions, so that good news can appear more credible and be better accepted by investors. We, therefore expect that readability of analyst reports has a stronger impact on upward revision reports, but not so when analysts issue downward revision forecasts.

Second, analysts' forecasts are greatly influenced by information provided by the management (Chen and Matsumoto, 2006). Managers are more likely to provide quality information when there is good news so that it will be better received by investors. Hutton et al. (2003) provided empirical evidence to the effect that management forecasts are more likely to be supported by verifiable soft talk when they issue good-news management forecasts, but not when they issue bad-news management forecasts. As a result, when analyst forecast revisions are driven by good news from management, analysts have better quality information, i.e. more precise information inputs, for more accurate forecasts and for better writing their reports and, therefore, evoke stronger market reactions.

Overall, we expect that readability indicates better information only when there is an upward revision. This prediction is consistent with the notion that analysts have an incentive to issue optimistic forecasts to please the management. Specifically, we expect that rational investors rely more on readable reports with upward revisions when forming their

expectations of future earnings and the market is more sensitive to upward revisions of earnings forecasts when the report is more readable.

We don't have a clear understanding of influence of readability on market reaction to downward revisions of forecasts. Given analysts' incentive to issue optimistic forecasts to please the management, and the market participants being less skeptical about the credibility of downward revisions, this mitigates the need to use readability to enhance the credibility of forecasts.⁵ Hypothesis 1 is developed as follows:

Hypothesis 1: Market reaction to forecast revisions is stronger when reports are more readable and the revisions are upward.

In the second part of the tests, we further confirm whether the prediction in the first part is due to the positive association between readability and accuracy of analyst forecasts, and test the relationship between readability and accuracy. Specifically, if the result of testing Hypothesis 1 indicates that the market does react more strongly to more readable reports, due to the expectation of better accuracy, we expect readability to be positively associated with accuracy, measured ex-post.

On the other hand, if analysts write readable reports regardless of information quality and accuracy, the prediction wouldn't hold. In extreme cases, we expect the opposite, i.e. if analysts write readable reports in order to bias market expectations, that may cause a negative association between readability and forecast accuracy. For example, an analyst who has poor information may write an easily understandable report to bias the reader and evoke stronger investor reaction to upward revisions or stock recommendations.

In addition, following our conjecture that readability is more important when there is an upward revision because then analysts are more likely to provide readable reports (so that

⁵ Prior studies suggest that analysts are more likely to issue optimistic forecasts and good news when they are affiliated with the firm (O'Brien et al., 2005). The opposite may be true if affiliated analysts issue readable reports to bias the expectation upwards, to please the firms.

the upward revisions can be better perceived by investors), we expect analysts to use readability to signal accuracy when there is an upward revision. As a result, we expect that readability is related to accuracy when there is an upward revision, but not in case of downward revision of expected earnings. Hypothesis 2:

Hypothesis 2: Readability of reports is associated with ex-post realized accuracy of earnings forecasts, and this association is attributed to forecast revisions being upward.

In the third part of this paper, we examine the determinants of report readability. Specifically, we study determinants of report readability in the context of analyst characteristics, information environment and characteristics of the covered firms. First, analyst characteristics indicate the ability to forecast accurately. Since more readable reports enable information to be better understood by investors, analysts who are able to forecast more accurately have stronger incentives to put in greater effort to make the report more readable, so that they can maximize (minimize) the benefit (cost) of accurate (inaccurate) forecasts. If this is the case, we would expect a positive association between report readability and analyst ability to issue accurate forecasts. Hypothesis 3A (null):

Hypothesis 3A: Readability of reports is associated with analyst ability to forecast future earnings accurately.

The second context is the information environment of the firm the analysts are following. Baginski and Hassell (1990), and Lehavy et al. (2011) show that accounting information helps analysts better predict future performance. Their papers provide empirical evidence that revisions issued immediately after management forecasts are more accurate than those issued before the announcements. Accounting information may also trigger

gathering of more private information. Barron, Byard and Kim (2002) provided evidence that private information acquisition by security analysts increases after quarterly earnings announcements. Greater private information also improves accuracy of forecasts. Overall, empirical evidence suggests that accounting disclosures improve the availability of firm specific information in the market, which enables analysts to issue more accurate forecasts and write more readable reports. As a result, we expect analyst reports issued concurrently with, or immediately after, accounting disclosures (i.e. earnings announcements or management forecasts) are more readable because of the better information available for making accurate forecasts. Hypothesis 3B (null):

Hypothesis 3B: Readability of reports is associated with whether analyst reports are issued concurrently with, or immediately after, management forecasts or earnings announcements.

In our empirical tests, as an alternative measure of information availability, we also examine the influence of firm characteristics on analyst reports' readability (e.g., Li, 2008). More information is required to predict future performance if a firm operates in a more complicated and uncertain business environment or sector. An analyst will, therefore, be less likely to write a readable report if the firm operates in multiple segments or is large in size. On the other hand, an analyst is more likely to write readable reports if the firm provides more information because of events such as mergers and acquisitions and has a long operating history. Hypothesis 3C (null):

Hypothesis 3C: Readability of reports is associated with characteristics of the firm that the report covers.

In the last part of the tests, we partition our sample into upward and downward revisions. We follow the same logic as in Hypothesis 1 and 2, i.e. analysts have stronger incentive to issue readable reports when revising forecasts upwards. We expect the effect of analysts characteristics on forecasting ability better explains report readability when there is upward revision. Hypothesis 3D (null):

Hypothesis 3D: The determinants of report readability are different for upward and downward revisions.

3. Research Plan and Methodology

3.1 Measures of analyst report readability

To measure readability, we adopt measures of readability of disclosures used in prior literature (Soper and Dolphin, 1964; Schroeder and Gibson, 1990; Li, 2008). We discuss the measures of readability in our tests as follows.

- 1) Fog index: This index indicates the number of years of education required to understand a text, with a lower value implying that the text is easier to understand. The variable Fog is measured as $0.4 * ((\text{words/sentence}) + 100 (\text{complex words/words}))$.
- 2) Flesch reading ease score: This variable is similar to Fog in meaning, but is constructed in such a way that higher scores mean greater readability. The variable Flesch is measured as $206.835 - 1.015 * (\text{total words/total sentences}) - 84.6 * (\text{total syllables/total words})$.
- 3) Flesch-Kincaid grade level: This formula translates scores into a U.S. grade level, or number of years of education required to understand the text. Similar to the Fog index, lower values indicate more readable texts. The variable Flesch_Kincaid is used to

represent this measure, and is calculated as $0.39 * (\text{total words}/\text{total sentences}) + 11.8 * (\text{total syllables} / \text{total words}) - 15.59$.

To formulate a readability measure, we calculate the mean value of standardized scores of Fog, Flesch and Flesch_Kincaid across all sample reports and denote it as *Totalscore*, and define a Standardized Fog score = $(S - \text{MAX})/(\text{MIN} - \text{MAX})$, where S = Fog score, MAX = maximum Fog score across reports and MIN = minimum Fog score across reports. The Standardized Flesch-Kincaid score is defined in the same way. The Standardized Flesch score is defined as $(S - \text{MIN})/(\text{MAX} - \text{MIN})$.

For better understanding of “readability” of analyst reports, Appendix A shows two analyst reports with different levels of readability. The examples are the summary sections of the beginning of two analyst reports. For comparability, we select two reports about the same company issued on the same date, but by two analysts. The first report has fewer words (155) than the second report (214). However, the first report contains more complex words with average 1.79 syllables as compared to 1.58 syllables in the second report. The first report also has average of 25.83 words per sentence as compared to 6.29 words per sentence in the second report. As a result, for each of our readability measure, report one is identified as the less readable report.

3.2 Measures of stock returns and investment decisions

Stock returns are measured across the following two windows. First, we measure the $[-1, +1]$ days window around the time of analyst report issuance. This event allows us to establish whether the return (adjusted for market return) differs across reports with different levels of readability.

Specifically, we test whether sensitivity of market reaction to forecast revisions differs with report readability, using the following regression.

$$CAR[-1, +1] = a + b*Totalscore + c*News + d*Totalscore*News + Controls + e, \quad (1)$$

where *Totalscore* is measured as previously explained. CAR is the market adjusted cumulative abnormal returns for [-1, +1] days around/following the forecast revision date. *News* is the difference between analyst EPS forecast on the date of issuance of the report (for the most recent period) and mean value of analysts' consensus forecast for the same firm, for the same period, scaled by the absolute value of the consensus forecast. Control variables include market value of equity (MVE), market to book ratio (MTB) and 180-days buy-and-hold return up to the report's date (Momentum). MVE is the log value of market value of equity. MTB is defined as market value of equity plus book value of liabilities scaled by total assets. Following Hypothesis 1, we expect *d* to be significantly positive in Equation (1), given the readability increases the expected accuracy of the forecast and, therefore, the forecast's response coefficient. Specifically, since analysts have stronger incentives to issue readable reports when there is an upward revision, we expect *d* to be significantly positive when the revision contains good news. We, therefore, construct a dichotomous variable (*Gnews*) to partition our sample. The variable equals one if the variable *News* is positive, and zero otherwise. To control for industry specific omitted variables, we include industry fixed effects (4-digit SIC) in all regressions.

3.3 Measures of ex-post forecast accuracy and biases

We test Hypotheses 2A and 2B in regression (2):

$$|FE| = a + b*Totalscore + Controls + e, \quad (2)$$

$|FE|$ is the absolute value of forecast error FE , where forecast error FE is defined as the difference between an analyst's forecast and the actual realized earnings for the same period, scaled by stock price at the previous fiscal quarter end.

We also include market value (MVE) and market to book ratio (MTB) to control for the information environment. In addition, we control for return volatility ($RET_volatility$) and earnings volatility ($EPS_volatility$) as measures of uncertainty of predicting future earnings. $RET_volatility$ is the standard deviation of 15-month stock returns, and $EPS_volatility$ is the standard deviation of EPS during the preceding 16 quarters. We further control for possible earnings management due to debt covenants using *Leverage*, defined as the ratio of total assets to stockholders equity for the most recent quarter. To further control for predictability of forecast earnings, we control for earnings growth ($Egrowth$) and whether the firm has loss ($PredictE_loss$). $Egrowth$ is earnings growth, measured as the change of income before extraordinary items for the most recent quarter, scaled by income before extraordinary items in the earlier quarter. $PredictE_loss$ is a dichotomous variable that equals one if the analyst predicted earnings are negative, and zero otherwise. Following Hypothesis 2, if readability reflects ex-post forecast accuracy, we expect b to be significantly positive in Equation (2). In addition, if readability is more important for forecast accuracy when there is upward revision, we expect coefficient d to be significantly positive in Equation (2) when the report contains upward revision.

3.4 Tests for determinants of readability

In this part of the study, we examine the association between linguistic characteristics of analyst reports and various other factors that determine readability, including analyst's characteristics (H3A), public information released by the firm (H3B), and the firm's characteristics (H3C). The following regression model is used to test Hypotheses 3A to 3C.

$$Totalscore = a + b*Analyst\ Factors + c*Report\ Factors + d*Firm\ Factors + e, \quad (3)$$

Where *Totalscore* is as defined earlier. The factors are defined as follows (see Appendix B for details). *Analyst Factors* is a vector of variables measuring the ability and experience of an analyst. These variables include analysts' general experience (*G_experience*) and firm-specific experience (*F_experience*), number of firms and industries covered (*NFirm* and *Nindustry*), forecast frequency (*Frequency*), broker size (*Broker_size*) and historical accuracy (*Accuracy*). Specifically, we expect positive coefficients of experience, size of the brokerage house and previous record of accuracy (Mikhail, Walther and Willis, 1997; Clement, 1999; Jacob, Lys and Neale, 1999). *G_experience* is the total number of years for which an analyst has supplied at least one forecast. *F_experience* is the number of years for which an analyst has supplied at least one forecast for a specific firm during a firm-year. *Broker_size* measures the number of analysts working for the broker that employs the analyst, and *Accuracy* is the analyst's historical forecast precision, measured as the inverse of the analyst's average absolute forecast errors ($|FE|$, defined as before) between 1986 and 2007. Since analysts who have issued more accurate forecasts in the past are more likely to issue accurate forecasts in the current year also, we expect the coefficient of *Accuracy* to be positive. On the other hand, we expect a negative correlation between *NFirm* (*NIndustry*) and readability since greater firm or industry coverage means less effort or attention when forecasting future performance of one firm or industry (Clement, 1999; Jacob et al., 1999). *NFirm* (*NIndustry*) measures the number of firms (industries) covered by the analyst. We expect that higher forecast frequency for a specific firm means better information and greater effort put in for forecast revision. As a result, the coefficient of *For_frequency* is expected to be positive. For *Report Factors*, we try to capture the characteristics of analyst reports in regard to the following aspects. First, we test whether the information from management influences readability. Specifically, we consider whether earnings announcements or management guidance has been issued

immediately before an analyst report (Baginski and Hassell, 1990). Specifically, *Release* (*Guide*) is a dichotomous variable that equals one if there is an earnings release (guidance) issued within five days before the analyst report, and zero otherwise. Following Hypothesis 2, we expect *Release* and *Guide* to be significantly positive if the information provided in the company's public disclosures improves analyst report readability. We expect coefficients of *Release* and *Guide* to be significantly positive since management guidance provides information to improve the ex-post accuracy of forecasts. Second, we conjecture that reports serve to explain analysts' final products, i.e. stock recommendations. We measure an analyst's recommendation adjusted by change of mean consensus recommendations (*REC*) and the measure takes a higher value when there is an upgrade. Following the logic that analysts have stronger incentives to enhance readability to signal accuracy of upward revisions, we expect readability is higher when the recommendation has an upgrade. Third, readability of a report may be affected by whether the report is completed by a single analyst or a team (Brown and Hugon, 2009). We, therefore, include the variable *Team*, measured as the number of analysts listed in a specific report. Given prior findings that analyst teams issue less accurate forecasts than single analysts, we expect analyst team has a negative impact on report readability. Furthermore, if more reports are issued on the same day, analysts are more likely to herd and share common information (Trueman, 1994; Hong, Kubik and Solomon, 2000). The variable *NReport* measures the number of analyst reports issued on the same day for the same firm.

For *Firm Factors*, we also include firm characteristics as other determinants of analyst report readability (Li, 2008). These variables include log value of market value of equity (*MVE*), market-to-book ratio (*MTB*), return volatility measured as standard deviation of 15-month stock returns (*RET_volatility*) and earnings volatility (*EPS_volatility*). Other variables include dichotomous variables that measure whether a firm engaged in mergers and

acquisitions or made secondary equity offerings in a specific quarter (*MAQR*; *SEOQR*), the number of business and geographic segments (*NBSEG*; *NGSEG*), logarithm of the number of non-missing items in Compustat (*NITEM*), special items scaled by total assets (*SI*), firm age (*Age*), and a dummy variable indicating whether a firm is incorporated in Delaware (*DLW*). We expect firms with more complicated operations, i.e. bigger size (*MVE*), high growth (*MTB*), more volatile performance (*RET_volatility* and *EPS_volatility*), greater number of non-missing items (*NITEM*) and special items (*SI*), greater number of business and geographic segments (*NBSEG*; *NGSEG*), and incorporated in Delaware (*DLW*) to be associated with less readable reports. On the other hand, we expect reports to be more readable when the market demands greater amount of information from the firm, i.e. bigger size and higher growth (*MVE* and *MTB*), the firm has made secondary equity offerings in a specific quarter (*MAQR*; *SEOQR*), or has better information environment, i.e. longer firm age (*AGE*).

3.5 Data

We use the *Investext* database provided by Thomson Financial for our analyst report data. This is the world's largest online database for company and industry research, and provides full texts of two million company, industry and geographic research reports written by analysts from more than 600 leading investment banks, brokerage houses and consulting firms worldwide. Investext also provides in-depth analyses and data on approximately 60,000 publicly traded companies, including sales and earnings forecasts, market projections and research and development expenditure.

The analyst reports were first downloaded from Investext in PDF format. To analyze the readability of these reports, we first converted the files into machine-readable form. An important step was to manually select discussion parts from each report instead of batch

transformation by computers. The reason was that similar to companies' annual reports but unlike 10-K filings, the typesetting of an analyst report is usually complicated and varies across brokerage houses and analysts. An analyst report typically contains brokerage firms' logos, figures and tables. In addition, there are texts other than company analysis, including disclaimers, analyst qualifications, analyst contacts (e.g., phone numbers and emails) and page numbers, etc. These irrelevant pieces of information may cause measurement errors that are possibly correlated with our testing variables. To reduce these measurement errors, we manually collected and verified each analyst report, before using computing software to calculate the readability index to formulate our report readability measures.

We started our analysis using data of the hi-tech industry, as this industry has generated great research interest from investors, given its high growth and high level of uncertainty in valuation. Prior studies (Barth, Kasznik and McNichols, 2002; and Barron, Byard, Kile and Riedl, 2002) provide evidence that investors demand more information from securities analysts when the firms have higher growth and a larger proportion of intangible assets. Firms with greater levels of intangible assets also attract greater analyst coverage and more analyst research (Barron et al., 2002). The hi-tech industry has higher growth, its market-to-book ratio is 30% higher than non-hi-tech industries (Appendix B), and it has more intangible assets (measured by R&D and advertising expenses), about 8.3 times non-hi-tech industry. The high-tech industry has larger cross-sectional variations in growth and intangible assets than other firms. Standard deviation of MTB ratio is 36% higher, and that of intangible assets is 3.29 times greater than other firms. Appendix B also shows that the hi-tech industry is followed by more analysts than other industries, with an average of 6.90 analysts following each firm. Other industries in Compustat have an average of 6.03 analysts following each firm, which is 14% less than the hi-tech industry. A high number of analysts following a firm helps generate a larger sample per firm and to perform more powerful tests across analysts

covering the same firm. As a result, focusing on the high-tech industry enhances the ability to test report readability, and provides important insights into analysts' role in valuation of the capital market.

Sample descriptions are reported in Table 1. We obtained a total of 2,930 reports for 132 firms for the years 2005-2007. Panel A shows the sample selection. Most of the sample has been deleted because of analyst names missing from the reports and/or missing analyst forecast data from I/B/E/S. The second part of Panel A shows distribution of the sample over industries. The majority of sample firms are from the software industry. Prepackaged Software (SIC=7372) accounts for a large percentage of reports in our sample (46.4%). The industry distribution across our sample is consistent with the distribution in Compustat data.

Panel B shows descriptive statistics of variables. At the report level, it requires, on average, around 14 years of education to understand an analyst report (mean *Fog* = 14.01), indicating that analyst reports may not be easily understood by investors without a college education. Analyst reports are usually completed by a team (mean *Team* = 0.70). The mean of *NReport* is 6.40, suggesting that analysts cluster in issuing reports for the same firm on the same day.⁶ Moreover, analyst reports are often issued immediately after earnings releases or issuance of management guidance, as mean values of *Release* and *Guide* are 70% and 75%, respectively. Analysts have an average of seven years of general experience and three years of firm-specific experience. While they tend to follow large numbers of firms, they mostly concentrate on only a few industries. As is typical for high-tech firms, firm level variables show that firms are mostly young (mean *Age* = 16.78), have high market-to-book ratio, and are concentrated in small numbers of business or geographical segments.

⁶ Note that this variable considers all reports issued, regardless of whether they have sufficient data to be included in our sample.

Table 2 shows variables correlations. Pearson correlations are reported in the lower triangle and Spearman correlations are reported in the upper triangle. The three measures of readability are highly correlated (over 88%). Panel B shows correlations of our variables. None of the variables have correlations high enough to raise the concern of multicollinearity. Variable *Totalscore* is significantly correlated with analysts' firm-specific experience (0.04 in Pearson correlations), number of firms or number of industries followed (0.10), and the tenor of forecast news (0.03). This suggests that readability is positively related to analyst experience. In addition, *Totalscore* is also significantly higher when reports are issued after earnings announcements (0.22) and guidance issuances (0.21). This suggests that reports issued right after company announcements have higher readability.

4. Empirical Findings

4.1 Market reactions

In Table 3, we use OLS regression to test market reactions to forecast revisions conditional on readability of reports. In Panel A, we study event window $[-1, +1]$ around the time of report announcements. In the full sample test (Column 1), we find that the coefficient of *Totalscore* is significantly positive ($t=3.53$), but *Totalscore*News* is insignificantly positive ($t=0.68$). This result suggests that the markets react more positively on the date of announcement when the report is more readable. However, the overall market reaction to forecast news is not influenced by report readability. To further test whether readable reports are more informative when the revision is upward, we segregate positive and negative forecast news. Column 2 shows the results for upward forecast revisions ($GNews = 1$). Consistent with Hypothesis 1, we find that the coefficient of *Totalscore* News* is significantly positive ($t=2.21$). On the other hand, we don't find similar results in Column 3 ($t=0.08$) when the report has downward revisions ($GNews = 0$). This result is consistent with

H1; analysts have stronger incentive to issue readable reports when making upward revisions, to improve market acceptance of the revision, and the market perceives readable upward revision reports as more precise and, therefore, assigns higher weights to forecasts when revising beliefs of firm value.

To further improve the confidence of our findings being consistent with readability reflecting analysts' willingness to promote quality forecasts, we provide further cross-sectional tests to study whether the coefficient of *Totalscore*News* varies across two dimensions: 1) their confidence of forecast accuracy (measured by the quality of forecasting information and expected forecast accuracy; 2) the reputation effect of forecasts (i.e. investors' demand of information and price impact of forecasts).

Table 3 Panel B provide results of tests conducted to determine whether readability has stronger impact on investment decisions when the analysts are better informed and confident of forecast accuracy. We perform the test by dividing the sample on the basis of forecast horizon. We expect that readable reports have greater influence on market reaction when forecasts cover shorter horizons. Given analysts have better information when the horizon is short, implying greater confidence in accuracy of their analyses, we expect them to choose to issue more readable reports when forecast horizon is short. In Columns 1 to 3, we use 1,525 observations with forecast horizons below the sample median. And in Columns 4 to 6, we report test results for reports with forecast horizons above the sample median. We further split the sample based on forecast news. We expect the coefficient of *Totalscore* News* to be significantly positive when the horizon is shorter than the sample median, but not when the horizon is longer. The results in Table 3 Panel B support our argument. We find a significant coefficient of *Totalscore* News* ($t=2.05$) when the horizon is shorter than the sample median, but not when the horizon is longer ($t=0.74$). This finding confirms our expectation that analysts' issue readable forecasts only when they are more certain about accuracy.

To further confirm the expectation that analysts issue readable reports when they have better confidence on accuracy. We also partition our tests based on ex-post forecast accuracy, and expect that the coefficient of *Totalscore* News* is significantly positive only when the ex-post forecast error is small. The results (un-tabulated) indicate that the coefficient of *Totalscore* News* is significantly positive ($t=1.88$) when the forecast error is below the sample median, but not when the errors are above the sample median ($t=-0.48$). Under this construct, we further examine whether analysts have incentive to issue readable reports when there is an upward revision. Consistent with the findings in Table 3 Panel B, the coefficient of *Totalscore* News* is significantly positive ($t=1.89$) when an upward revision report has a forecast horizon below the sample median. The coefficient is not significant when the forecast horizon is longer or there is a downward revision in the report.

In Panel D of Table 3, we partition the sample using bid-ask spread as a measure of information asymmetry in the market. Since market has stronger demand for information to interpret analyst forecasts when information asymmetry is higher, and since readability indicates accuracy, we expect readability to have greater impact on market belief when the bid-ask spread is wider. To perform the test, we first divide the sample according to whether the bid-ask spread is greater than the sample median, and then divide the sample into upward and downward revisions. Consistent with our expectation, the results in Table 3 Panel C suggest that the coefficient of *Totalscore*News* is significantly positive when $GNews=1$ and bid-ask spread is greater than the sample median ($t\text{-stat} = 2.43$). Coefficients are insignificant in other cases.

In Panel E of Table 3, we partition the sample by identifying whether the forecast is a high/low innovative forecast. Gleason and Lee (2003) found that high-innovation forecast revisions (new forecast above the analyst's prior forecast and the consensus forecast, or below both forecasts) have a greater price effect than low-innovation forecasts. As a result,

we expect analysts to be more concerned about readability in case of high innovation forecast than low innovation forecasts. This result is consistent with our expectation; the coefficient of *Totalscore*News* is significantly more positive (t-statistic = 2.01) only in the sub-sample of upward revisions and high innovative forecast revisions.

4.2 Ex-post accuracy

Table 4 reports results in terms of ex-post forecast accuracy. Specifically, we use absolute forecast errors $|FE|$ as the dependent variable and test whether readability as measured by *Totalscore* is negatively associated with forecast errors. OLS multi-variant tests are presented. Column 1 reports the full-sample results. The coefficient of *Totalscore* is not significant (t=0.55), consistent with the insignificant findings in Table 3 regarding the relationship between market reactions to forecasts not being significantly different across reports. Similar to Table 3, we further separate our sample into two groups based on forecast news. In Column 2 ($GNews = 1$), the coefficient of report readability (*Totalscore*) is significantly negative (t=-1.74) at two-tailed 10% level. This suggests that more readable reports have smaller ex-post forecast errors if the reports have upward revision news. This finding is consistent with Hypothesis 2, that is, readability is indicative of forecast accuracy when the report has an upward revision. On the contrary, the coefficient of *Totalscore* is significantly positive in Column 3 ($GNews = 0$, t=1.92), suggesting that when the revision is downward, report readability is associated with greater forecast errors. This finding suggests that when managers issue less readable reports, downward revisions are more accurate ex-post. Overall, consistent with Hypothesis 2, Table 4 suggests that analysts issue more accurate and readable reports when there is upward revision.⁷

⁷ On the bad news side, analysts issue less accurate forecasts when the report is readable. This may be consistent with analysts' incentive to issue less readable reports when the bad news is more accurate in order to please the management. Consistent with this conjecture, the insignificant coefficient of readability and revision news in Table 3 also suggests that the market may partially see through this bias, and on average reacts little to

4.3 Determinants of Readability

4.3.1 Analyst Factors

Table 5 reports the results of determinants tests. The dependent variable is the readability score *Totalscore*. We divided the testing variables into three groups. Following H3A, the first group is to test whether *Analyst Factors* affect report readability. Our results show that firm experience *F_experience* is significantly positive at 1% ($t=3.35$). This suggests that analysts with longer experience of following a specific firm are more capable of issuing readable reports. Effect of the number of firms followed by an analyst, *NFirm*, is also significantly positive ($t=4.62$). This result suggests that firm coverage is positively associated with readability. We also find a marginally significant result for forecast frequency *For_frequency* ($t=1.75$). Firm experience represents an analyst's firm specific knowledge for forecasting future earnings (Mikhail et al., 1999). Prior literature suggests that analysts with longer experience (Mikhail et al., 1997) and prior record of accuracy (Hong and Kubik, 2003) have greater ability to provide accurate forecasts. The positive associations between firm experience and readability imply that analysts with better capabilities and those who put in greater effort in forecasting future earnings produce reports that are more readable. However, we find that larger number of firms covered being positively associated with readability is opposite to our expectation, given that a larger number of firms covered may divert analyst efforts to follow an individual firm. One possible explanation could be that in accordance with literature on industry information transfer, analysts mostly cover firms in the same industry and, therefore, covering other similar firms improves analysts' information set and their ability to issue accurate forecasts (Hui and Young, 2011). Overall, results in this section

provide mixed evidence of readability of analyst reports being associated with analysts who can issue more accurate forecasts.

In the second and third columns of Table 5 Panel A, we divide our sample based on the type of report news. Specifically, we test the difference between coefficients across the two columns using Chi-square test (untabulated). The results suggest that smaller brokerage houses (*Broker_size*, chi-square=4.07) and analysts with better prior records of forecast accuracy (*Accuracy*, chi-square=6.79) are significantly more likely to issue readable upward revision reports. Forecast accuracy is derived from verifiable records of analysts' abilities to predict future earnings. Smaller forecast errors mean more capable analysts. Consistent with Hypothesis 1 and 2, the significantly positive coefficient in upward revision sub-sample (Column 2) suggests that analysts issue readable reports when they are more capable of issuing accurate upward forecast revisions. This finding is consistent with the notion that readability enhances the credibility of upward revisions, but not of downward revisions. The brokerage-size result is consistent with smaller brokerage houses being concerned about their relationship with the management for better information inputs (Ke and Yu, 2006) and, therefore, being more likely to improve the readability of good-news reports. This result is also consistent with the theory that as analysts in early phases of their careers (when working for smaller brokerage houses) are more concerned about the reputation of their forecasts and are, therefore, more willing to exert effort to improve readability to enhance the quality of their forecasts (Holmstrom, 1999).

4.3.2 Report Factors

The second part of Table 5 tests H3B. We present the results of whether report characteristics (i.e. *Report Factors*) determine readability. Specifically, we find that analyst reports issued immediately after earnings announcements and/or management forecasts are more readable. Coefficients of *Release* and *Guide* are significantly positive at 1% level.

Release and *Guide* are variables indicating whether analyst reports are issued immediately after management information releases, i.e. earnings announcements and management forecasts. Managerial information is superior to analyst forecasts before public disclosure (Baginski and Hassell, 1990). The evidence that management information improves readability, therefore, implies that readability is a positive function of the underlying information set processed by analysts since better information set is positively associated with ex-post forecast accuracy. This finding confirms Hypothesis 3C, that is, firm disclosures influence analyst report readability, and also supports our expectation that analysts with ability to make more accurate forecasts are willing to issue more accurate forecasts.

Other report characteristics are also associated with readability. Specifically, *REC* is significantly positive. This suggests that consistent with our expectations, analysts issue more readable reports when there is a recommendation upgrade. *Team* is significantly negative, which means greater news surprises in the revision and team work result in less readable reports. This is consistent with our expectation that team work leads to less accurate forecasts (Brown and Hugon, 2009) and, therefore, lower analysts' incentive to improve readability. We also control for report length in our tests (*lnText_length*). By construction of the variables, *lnText_length* is significantly negative. We use report length to control for information in the report that is not controllable by analysts. For example, a firm may have great earnings surprises in a specific quarter, and this large surprise naturally causes the report to be longer than usual.

In the second and third columns, we don't find earnings announcements (*Release*) resulting in different results for upward and downward revision reports. This suggests that analyst report readability is equally driven by earnings announcements, regardless of the direction of report news. On the other hand, we find that results for *Guide* are significantly more positive (chi-square=2.89) at 10% level, when the report has downward revision news.

This result is consistent with prior literature in that market participants perceive bad-news management forecasts as more informative than good-news management forecasts (Skinner, 1994). As a result, analysts are more willing to write readable reports following bad-news management forecasts due to its informativeness in predicting future earnings (e.g., Skinner, 1994).

4.3.3 Firm Factors

In the last part of Table 5, we include firm characteristics in our tests (i.e. *Firm Factors*). Firms with higher growth (*MTB*) ($t=4.78$), fewer mergers and acquisitions (*MAQR*) ($t=-3.33$), longer history (*AGE*) ($t=2.25$) and incorporated outside Delaware (*DLW*) ($t=-2.56$) have more readable reports. Investors have greater demand for services of securities analysts when firms are smaller, and have high growth prospects and complex operations. As a result, analysts may be more willing to provide readable reports to satisfy the demand for information. This result is also consistent with Hypothesis 3C; firm characteristics affect analyst report readability. In the second and third columns, we find that earnings volatility (*EPS_volatility*, chi-square=4.24) and the number of segments (*SEOQR*, chi-square=68.15) have significantly stronger impact on readability of bad news analyst reports than upward revision reports.

Overall, results in second and third columns of Table 5 suggest that upward and downward revision reports are systemically different and are driven by different factors. Downward revision reports' readability is more sensitive to information environment and earnings predictability as measured by earnings volatility, firm age and management guidance. Upward revision reports' readability is more sensitive to analysts and broker characteristics, e.g., analyst forecast accuracy and broker size.⁸

⁸ Our results are robust after controlling for the tone of analyst report.

5. Conclusion

In this paper, we study readability of analyst reports. In the first and second parts of the paper, we examine the market's interpretation of analyst report readability and whether readability predicts ex-post forecast errors. The results suggest that readability of upward revision reports indicates smaller forecast errors. Market participants also understand this and, therefore, have a higher forecast earnings response coefficient when an upward revision report is more readable. Moreover, they are indifferent to readability in downward revision reports when forming their expectations. In the third part of the tests, we find that analysts with better experience, who follow more firms and have a better forecast accuracy history, write more readable reports. We also find that analyst reports are more readable when they are issued immediately after public disclosures (i.e. earnings announcements or management forecasts). This result suggests that readability of analyst report is greatly influenced by analyst characteristics, firms' disclosure policies and other firm characteristics. We also divide our tests into sub-samples of firms that issue upward and downward revision reports. We find that analysts with better prior forecast accuracy and those from smaller brokerage houses write more readable upward revision reports, but not more readable downward revision reports.

This paper contributes to financial analysis literature by studying linguistic characteristics of information publicly available to investors but poorly understood by researchers. First, by examining whether the way in which reports are written affects the impact of information they convey. We find that both analyst characteristics and management influence the writing of research reports. Second, we extend the existing research on market reactions to analyst reports beyond earnings forecasts and stock recommendations. We show that readability of an analyst report can affect an investor's perception of the analyst's forecasts and recommendations and, therefore, lead to a better understanding of behaviors of

market participants. Third, we incorporate linguistic research (text readability) into analyst research. We explore determinants of the language that analysts use in their reports and examine whether these linguistic features are associated with the underlying characteristics of the analysts, their reports and the firms. This provides a different perspective of tone and content and contributes to the linguistic literature, as well as the literature on financial analysts.

Appendix A: Examples of Analyst Report Readability

Example 1: Analyst Report on Citrix Systems Inc, July 20, 2006

“CTXS posted a solid quarter, although unlike recent quarters in which they were firing on all cylinders CTXS in 2Q06 suffered some blemishes which weren't bad but which didn't make for the kind of out-performance the stock's premium valuation seemed to be discounting.

Higher costs absorbed higher revenues leading to inline EPS and guidance.

We don't think the quarter reflected a marked change in CTXS' business, though it could suggest that the positive benefit to license revs from end-of-life programs is largely over. It highlights though a willingness by CTXS to sacrifice near term profitability (or upside) for faster future growth that investors likely haven't fully appreciated.

We do believe the company will pursue additional investments/acquisitions which could prove dilutive, with WAN optimization companies topping the list, though CTXS' past success should give them better credibility with investors.

Near term though the stock is likely to remain under pressure as 'whisper expectations' of higher than consensus estimates get reigned in.”

Readability Indices

Number of words	: 155
Average syllables per word	: 1.79
Average words per sentence	: 25.83
Fog	: 17.82
Flesch	: 29.43
Flesch-Kincaid	: 15.57

EXAMPLE 2: Analyst Report on Citrix Systems Inc, July 20, 2006

“CTXS reported strong revenue, but increased expenses led to inline EPS. We remain buyers into what we believe will be weakness from these results. We decrease our price target from \$45 to \$42.

Revenue of \$275.5M exceeded our and consensus estimate of \$266M and above the range of guidance of \$254m to \$266m. EPS of \$0.33 was inline with consensus and at the top end of guidance. Expenses were higher due to an increase in headcount additions and the Reflectent acquisition, which closed in the quarter and was mildly dilutive.

Guidance was above street estimates, but only modestly above for EPS.

Guidance for Q3 was for revenue between \$275m and \$280m, with EPS between \$0.33 and \$0.34.

Guidance for FY 2006 was for revenue between \$1.105B and \$1.12B with EPS between \$1.36 and \$1.39.

Consensus revenue and EPS estimates for Q3 had been \$273m and \$0.34, and for FY 2006 \$1.1B and \$1.38.

We are making revisions to our estimates. For Q3, we increase our revenue estimate from \$272m to \$281m, and maintain our EPS estimate of \$0.34. For CY 2006, we are increasing our revenue estimate from \$1.09B to \$1.11B and increase our EPS estimate from \$1.38 to \$1.39. For CY 2007, we are increasing our \$1.22B revenue estimate to \$1.23B and decreasing our \$1.55 EPS estimate to \$1.54.

We are reiterating our Buy rating but lowering our price target from \$45 to \$42, which reflects a 27x multiple on C'07 EPS (a slight premium to the group but a decrease from our previous 29 multiple).”

Readability Indices

Number of words	: 214
Average syllables per word	: 1.58
Average words per sentence	: 6.29
Fog	: 9.81
Flesch	: 66.83
Flesch-Kincaid	: 5.50

Appendix B. Comparison between characteristics of High-Tech and Other Industries

The following table reports characteristics of high-tech industries, compared to those of other industries. “High-tech” industries are defined as in Baginski et al. (2004) (i.e. SIC codes = 2833-2836, 8731-8734, 7371-7379, 3570-3577, and 3600-3674). The “Other” industries are all remaining industries in Compustat. Analyst following data is obtained from I/B/E/S and accounting data from Compustat for fiscal years 2005-2007. All firm-level variables are winsorized for the top and bottom 1% of observations. Descriptive statistics for the two groups are reported as follows. *NUMEST* is the number of estimates for annual earnings. *lnTA* is the log of total assets. *MTB* is market -to-book ratio, measured as market value of equity divided by common stock at fiscal year end. *ROA* is return on assets, measured as income before extraordinary items divided by total assets. *Intangibles* are research and development expenses plus advertising expenses, divided by total operating expenses. T values (Z values) are reported for t-test (Wilcoxon two-sample test) for differences in means (medians). ***, **, and * represent significance at 1%, 5% and 10% levels, respectively.

Comparison across Compustat Industries based on Industry Data for Fiscal Year 2005-2007

	N		Mean		1 st Quartile		Median		3 rd Quartile		Std.		t-test for means (t)		Wilcoxon test for medians (Z)	
	High-tech	Other	High-tech	Other	High-tech	Other	High-tech	Other	High-tech	Other	High-tech	Other				
NUMEST	1,905	7,306	6.90	6.03	2.08	2.00	4.83	4.25	9.56	8.42	6.20	5.37	5.6	***	4.5	***
lnTA	1,905	7,306	5.70	7.13	4.36	5.81	5.37	7.02	6.86	8.33	1.94	1.99	-28.1	***	-27.9	***
Sales/TA	1,905	7,306	0.70	0.88	0.39	0.26	0.64	0.71	0.92	1.25	0.49	0.77	-12.1	***	-5.0	***
MTB	1,905	7,306	3.74	2.86	1.66	1.40	2.74	2.06	4.59	3.29	4.91	3.60	7.3	***	12.8	***
ROA	1,905	7,306	-0.04	0.02	-0.09	0.00	0.01	0.02	0.04	0.04	0.14	0.06	-17.3	***	-12.3	***
Intangibles	1,905	7,306	0.25	0.03	0.08	0.00	0.19	0.00	0.34	0.03	0.23	0.07	42.1	***	53.2	***

Variable	Description
<i>Report-specific</i>	
lnText_length	The natural log of number of words in the report.
Fog	Index of ease of reading. The index is lower if the text is easier to understand.
Flesch	Reading ease score. The index is higher if the text is easier to understand.
Flesch_Kincaid	Grade level for reading ease. The index is lower if the text is easier to understand.
Totalscore	Mean values of standardized scores of Fog, Flesch, and Flesch_Kincaid across all sample reports. Standardized Fog score = $(S - \text{MAX}) / (\text{MIN} - \text{MAX})$, where S = Fog score, MAX = maximum Fog score across reports, and MIN = minimum Fog score across reports. Standardized Flesch-Kincaid score is defined in the same way. Standardized Flesch score is defined as $(S - \text{MIN}) / (\text{MAX} - \text{MIN})$.
REC	The difference between an analyst's recommendation on the date of issuance of the report and mean value of analysts' consensus recommendation for the same firm, scaled by the consensus recommendation and multiplied by -1. Recommendation codes follow those used in I/B/E/S (i.e. 1 = strong buy; 2 = buy; 3 = hold; 4 = underperform; 5 = sell.). Original recommendations are recorded from sample reports and then cross-referenced with recommendation text from First Call. If a recommendation is missing in a report, it is replaced by the most recent recommendation in First Call issued by the same analyst.
News	The difference between the analyst's EPS forecast on the report date (for the most recent period) and mean value of analysts' consensus forecast for the same firm and the same period, scaled by the absolute value of the consensus forecast.
Gnews	A dichotomous variable that equals one if the variable <i>News</i> > 0; zero otherwise.
Team	A dichotomous variable that equals one if a specific report lists more than one analyst, and zero otherwise.
Nreport	Number of analyst reports issued on the same day for the same firm (including those whose analysts cannot be identified in I/B/E/S.)
Release	A dichotomous variable that equals one if there is an earnings release issued within five days before the analyst report date; zero otherwise.
Guide	A dichotomous variable that equals one if there is an earnings guidance issued within five days before the analyst report date; zero otherwise.
FE	The absolute value of ex-post forecast error where forecast error is defined as the difference between the analyst's forecast issued on the report date and the actual realized earnings for the forecast period, scaled by the most recent stock price, at fiscal quarter end. (Value is multiplied by 1,000.)
CAR[-1, +1]	Cumulative abnormal return for day -1 to day +1, where day 0 is the analyst report date. The estimation is based on market model using returns from day -131 to day -31.
Momentum	180-day buy-and-hold return up to a report's date.
<i>Analyst-specific</i>	
G_experience	Number of years for which analyst <i>i</i> supplied at least one forecast.
F_experience	Number of years through year <i>t</i> for which analyst <i>i</i> supplied at least one forecast.
Nfirm	A measure of the number of companies analyst <i>i</i> followed in year <i>t</i> .
Nindustry	A measure of the number of industries analyst <i>i</i> followed in year <i>t</i> .
For_frequency	Analyst <i>i</i> 's forecast frequency for firm <i>j</i> .
Broker_size	Number of analysts in the brokerage firm the analyst works for.
Accuracy	Analyst <i>i</i> 's historical forecast accuracy, measured as standardized ranking of the analyst's average absolute forecast errors (FE) between 1986 and

	2007, in each firm; FE is scaled by price.
<i>Firm-quarter-specific</i>	<i>(For the most recent quarter prior to the analyst report date)</i>
MVE	Logarithm of market value of equity.
MTB	Market value of equity plus book value of liabilities scaled by total assets.
RET_volatility	Standard deviation of 15-month stock returns for a specific quarter. PL CHECK THIS.
EPS_volatility	Standard deviation of EPS during the preceding 16 quarters.
Leverage	The ratio of total assets to stockholders equity for the most recent quarter.
Egrowth	The change of income before extraordinary items for the most recent quarter, scaled by income before extraordinary items in the preceding quarter.
PredictE_loss	A dichotomous variable that equals one if the analyst predicted earnings are negative, and zero otherwise
MAQR	A dichotomous variable that equals one if the firm appears in the SDC M&A database as an acquirer in the quarter; zero otherwise.
SEOQR	A dichotomous variable that equals one if the firm appears in the SDC database for issuing common equity in the quarter; zero otherwise.
NBSEG	Number of business segments.
NGSEG	Number of geographic segments.
NITEM	Logarithm of the number of non-missing items in Compustat.
SI	Special items scaled by total assets.
Age	Firm age since the firm first appeared in CRSP monthly stock file.
DLW	A dichotomous variable that equals one if a firm was incorporated at Delaware, and zero otherwise

REFERENCES

This is the Pre-Published Version

- Asquith, Paul, Michael B. Mikhail, and Andrea S. Au, 2005, Information content of equity analyst reports, *Journal of Financial Economics* 75, 245-282.
- Baginski, Stephen P., and John M. Hassell, 1990, The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revision, *The Accounting Review* 65, 175-190.
- Baginski, Stephen P., John M. Hassell and Michael D. Kimbrough, 2004, Why do managers explain their earnings forecasts, *Journal of Accounting Research* 42, 1-29.
- Barber, Brad M., Reuven Lehavy, Maureen McNichols, and Brett Trueman, 2003, Reassessing the returns to analysts' stock recommendations, *Financial Analysts Journal* 59, 88-96.
- Barber, Brad M., Reuven Lehavy, and Brett Trueman, 2007, Comparing the stock recommendation performance of investment banks and independent research firms, *Journal of Financial Economics* 85, 490-517.
- Barron, Orie E., Donal Byard, and Oliver Kim, 2002, Changes in analysts' information around earnings announcements, *The Accounting Review* 77, 821-846.
- Barron, Orie E., Donal Byard, C. Kile and E. Riedl, 2002, High-Technology intangibles and analysts' forecasts, *Journal of Accounting Research* 40, 289-312.
- Barth, M.m R. Kasznik and M. McNichols. Analyst coverage and intangible assets. *Journal of Accounting Research* 39: 1-34.
- Bloomfield, R. J. 2002. The "incomplete revelation hypothesis" and financial reporting. *Accounting Horizons* 16: 233-243.
- Brown, Lawrence D., and Artur Hugon, 2009, Team earnings forecasting, *Review of Accounting Studies* 14, 587-607.
- Clement, Michael B., 1999, Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27, 285-303.
- Chen, Shuping, and Dawn A. Matsumoto, 2006, Favorable versus unfavorable recommendations: The impact on analyst access to management-provided information, *Journal of Accounting Research* 44, 657-689.
- Clement, Michael B., and Senyo Y. Tse, 2003, Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78, 227-249.
- Cliff, Michael T., 2007, Do affiliated analysts mean what they say? *Financial Management* 36, 5-29.
- Francis, Jennifer, and Donna Philbrick, 1993, Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research* 31, 216-230.
- Gleason, Cristi A., and Charles M. C. Lee, 2003, Analyst forecast revisions and market price discovery, *The Accounting Review* 78, 193-225.
- Healy, Paul M., and Krishna G. Palepu, 2001, Information asymmetry, corporate disclosure and the capital markets: A review of the empirical disclosure literature, *Journal of Accounting and Economics* 31, 405-440.
- Holmstrom B., 1999, Managerial Incentive Problems: A Dynamic Perspective, *Review of Economic Studies*, 66, 169-182.
- Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon, 2000, Security analysts' career concerns and herding of earnings forecasts, *The Rand Journal of Economics* 31, 121-144.
- Hong, Harrison, and Jeffrey D. Kubik, 2003, Analyzing the analysts: Career concerns and biased earnings forecasts, *Journal of Finance* 58, 313-351.

- Hui, K. W., and P. Young, 2011, The post forecast revision drift and underreaction to industry-wide and/or firm-specific earnings, Working paper, Hong Kong University of Science and Technology.
- Hutton, Amy P., Gregory S. Miller, and Douglas J. Skinner, 2003, The role of supplementary statements with management earnings forecasts, *Journal of Accounting Research* 41, 867-890.
- Jacob, J., T. Lys, and M. Neale, 1999, Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (1): 27-50.
- Ke, Bin, and Yong Yu, 2006, The effect of issuing biased earnings forecasts on analysts' access to management and survival, *Journal of Accounting Research* 44, 965-999.
- Kothari, S. P., Xu Li, and James E. Short, 2009, The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis, *The Accounting Review* 84, 1639-1670.
- Li, Feng, 2008, Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics* 45, 221-247.
- Lehavy, R., Feng Li, K. Merkley, 2011, The effect of annual report readability on analyst following and the properties of their earnings forecasts, *The Accounting Review* 86, 1087-1106.
- Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis, 1997, Do security analysts improve their performance with experience? *Journal of Accounting Research* 35, 131-157.
- Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis, 1999, Does forecast accuracy matter to security analysts? *The Accounting Review* 74, 185-200.
- Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis, 2003, The effect of experience on security analyst underreaction, *Journal of Accounting and Economics* 35, 101-116.
- Michael B Mikhail, Beverly R Walther, Richard H Willis. 2007, When Security Analysts Talk, Who Listens? *The Accounting Review* 82, 1227-1254.
- O'Brien, Patricia C., Maureen F. McNichols, and Hsiou-Wei Lin, 2005, Analyst impartiality and investment banking relationships, *Journal of Accounting Research* 43, 623-650.
- Schroeder, Nicholas, and Charles Gibson, 1990, Readability of management's discussion and analysis, *Accounting Horizons* 4, 78-87.
- Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news, *Journal of Accounting Research* 32, 38-60.
- Soper, Fred J., and Robert Dolphin, Jr., 1964, Readability and corporate annual reports, *The Accounting Review* 39, 358-362.
- Stickel, Scott E., 1992, Reputation and performance among security analysts, *Journal of Finance* 47, 1811-1836.
- Trueman, Brett, 1994, Analyst forecasts and herding behavior, *The Review of Financial Studies* (1986-1998) 7, 97-124.

Table 1. Sample Distribution and Descriptive Statistics

This is the Pre-Published Version

Panel A. Sample selection

	N
Reports from Investext covering high-tech industry companies that have data in I/B/E/S on the same date for the same analyst during 2005 ~ 2007	3,554
Deduct:	
Reports without sufficient forecast and recommendation data in I/B/E/S and First Call	(435)
Reports without sufficient data in Compustat and CRSP	(189)
Sample number of reports:	<u>2,930</u>

Industry distribution

SIC	Industry name	Report level		Firm level	
		N	(%)	N	(%)
2834	Pharmaceutical Preparations	201	6.9	15	11.4
2835	In Vitro, In Vivo Diagnostics	10	0.3	1	0.8
2836	Biological Products, Except Diagnostic Substances	41	1.4	4	3.0
3570	Computer and Office Equipment	42	1.4	1	0.8
3571	Electronic Computers	73	2.5	1	0.8
3572	Electronic Computers	128	4.4	4	3.0
3576	Computer Communications Equipment	58	2.0	4	3.0
3577	Computer Peripheral Equipment, Not Elsewhere Classified	70	2.4	4	3.0
3600	Electronic and Other Electrical Equipment, Except Computer	3	0.1	1	0.8
3620	Electrical Industrial Apparatus	6	0.2	1	0.8
3630	Household Appliances	8	0.3	1	0.8
3640	Electric Lighting, Wiring Equipment	9	0.3	1	0.8
3661	Tele and Telegraph Apparatus	49	1.7	3	2.3
3663	Radio, TV Broadcast, Communication Equipment	179	6.1	9	6.8
3672	Printed Circuit Boards	11	0.4	3	2.3
3674	Semiconductor, Related Device	89	3.0	11	8.3
7371	Computer Programming Service	45	1.5	3	2.3
7372	Prepackaged Software	1,358	46.4	40	30.3
7373	Computer Integrated Systems Design	303	10.3	14	10.6
7374	Computer Processing, Data Preparation and Processing Services	182	6.2	7	5.3
8731	Commercial Physical and Biological Research	65	2.2	4	3.0
Total		<u>2,930</u>	<u>100</u>	<u>132</u>	<u>100</u>

Panel B. Descriptive statistics

	N	Mean	1st Quartile	Median	3rd Quartile	Std.
<i>Report-specific</i>						
lnText_length	2,930	6.72	6.28	6.79	7.14	0.64
Fog	2,930	14.01	12.53	13.82	15.32	2.08
Flesch	2,930	41.86	36.49	42.22	47.45	8.13
Flesch_Kincaid	2,930	10.67	9.33	10.48	11.87	1.89
Totalscore	2,930	55.84	42.98	57.08	69.78	19.52
REC	2,930	0.14	-0.24	-0.04	0.24	0.56
News	2,930	-0.04	-0.06	0.00	0.04	0.20
GNews	2,930	0.42	0.00	0.00	1.00	0.49
Team	2,930	0.70	0.00	1.00	1.00	0.46
NReport	2,930	6.40	2.00	6.00	10.00	4.45
Release	2,930	0.70	0.00	1.00	1.00	0.46
Guide	2,930	0.75	1.00	1.00	1.00	0.43
FE	2,930	1.81	0.43	0.97	2.22	2.50
CAR[-1, +1]	2,930	0.002	-0.038	0.003	0.049	0.080
<i>Analyst-specific</i>						
G_experience	2,930	7.08	4.00	7.00	8.00	4.19
F_experience	2,930	3.06	1.00	2.00	4.00	2.37
NFirm	2,930	18.11	14.00	18.00	22.00	6.17
NIndustry	2,930	2.97	2.00	3.00	4.00	1.31
For_frequency	2,930	9.95	5.00	9.00	13.00	5.72
Broker_size	2,930	94.03	37.00	73.00	160.00	71.69
Accuracy	2,930	0.50	0.33	0.51	0.66	0.22
<i>Firm-quarter-specific</i>						
MVE	2,930	8.45	6.97	8.38	9.66	1.78
MTB	2,930	2.93	1.91	2.50	3.57	1.53
RET_volatility	2,930	0.09	0.07	0.09	0.11	0.03
EPS_volatility	2,930	0.20	0.06	0.11	0.19	0.36
MAQR	2,930	0.25	0.00	0.00	0.00	0.43
SEOQR	2,930	0.00	0.00	0.00	0.00	0.04
NBSEG	2,930	1.09	0.69	0.69	1.61	0.53
NGSEG	2,930	1.48	1.39	1.39	1.79	0.40
NITEM	2,930	5.12	5.09	5.14	5.20	0.14
SI	2,930	-0.001	-0.001	0.000	0.000	0.004
Age	2,930	16.78	7.83	12.67	20.58	13.90
DLW	2,930	0.68	0.00	1.00	1.00	0.47

Note: lnText_length is the natural log of number of words in the report. Fog, Flesch, and Flesch_Kincaid are measures of readability. Totalscore is the mean value of standardized scores of Fog, Flesch, and Flesch_Kincaid across all sample reports. REC is the difference between the analyst's recommendation on the report date and mean value of analysts' consensus recommendation for the same firm, scaled by the consensus recommendation and multiplied by -1. Recommendation codes follow those used in I/B/E/S (i.e. 1 = strong buy; 2 = buy; 3 = hold; 4 = underperform; 5 = sell.). News is the difference between the analyst's EPS forecast issued on the report date (for the most recent period) and mean value of analysts' consensus forecast for the same firm and the same period, scaled by the absolute value of the consensus forecast. GNews is a dichotomous variable that equals one if the variable News > 0; zero otherwise. Team is a dichotomous variable that equals one if a specific report lists more than one analyst, and zero otherwise. NReport is the number of analyst reports issued on the same day for the same firm (including those whose analysts cannot be identified in I/B/E/S.) Release is a dichotomous variable that equals one if there is an earnings release issued within five days before the analyst report date; zero otherwise. Guide is a dichotomous variable that equals one if there is an earnings guidance issued within five days before the analyst report date; zero otherwise. |FE| is the absolute value of ex-post forecast error, where the forecast error is defined as the difference between the analyst's forecast on the report date and the actual realized earnings for the same period, scaled by the most recent stock price, at fiscal

quarter end. $CAR[-1, +1]$ is the cumulative abnormal return during day -1 to day +1, where day 0 is the analyst report date. The estimation is based on market model using returns during day -131 and day -31. $G_experience$ is the number of years for which analyst i supplied at least one forecast during year t . $F_experience$ is the number of years through year t for which analyst i supplied at least one forecast during year t for firm j . N_{firm} is the number of companies analyst i followed in year t . $N_{industry}$ is the number of industries analyst i followed in year t . $For_frequency$ is analyst i 's forecast frequency for firm j . $Broker_size$ is the number of analysts in the broker employing the analyst. $Accuracy$ is analyst i 's historical forecast accuracy, measured as the standardized ranking of the analyst's average absolute forecast errors ($|FE|$) during 1986 and 2007 for each firm. MVE is the logarithm of market value of equity. MTB is market value of equity plus book value of liability scaled by total assets. $RET_volatility$ is the standard deviation of 15-month stock returns for a specific quarter. $EPS_volatility$ is the standard deviation of EPS during the preceding 16 quarters. $MAQR$ is a dichotomous variable that equals one if the firm appears in the SDC M&A database as an acquirer in the quarter; zero otherwise. $SEOQR$ is a dichotomous variable that equals one if the firm appears in the SDC database for issuing common equity in the quarter; zero otherwise. $NBSEG$ is the number of business segments. $NGSEG$ is the number of geographic segments. $NITEM$ is the logarithm of the number of non-missing items in Compustat. SI is special items scaled by total assets. Age is firm age since the firm first appeared in CRSP monthly stock file. DLW is a dichotomous variable that equals one if a firm was incorporated at Delaware, and zero otherwise. See Appendix B for details.

Table 2. Variable Correlations^a**Panel A. Readability measures^b**

Variable	InText_length	Fog	Flesch	Flesch_Kincaid	Totalscore
InText_length		0.28	-0.25	0.33	-0.29
Fog	0.30		-0.88	0.95	-0.97
Flesch	-0.28	-0.89		-0.90	0.96
Flesch_Kincaid	0.35	0.96	-0.90		-0.98
Totalscore	-0.32	-0.98	0.96	-0.98	

Panel B. Univariate correlation between readability and testing variables^b

Variable	Totalscore	G_experience	F_experience	NFirm	NIndustry	For_frequency	Broker_size	Accuracy	Release	Guide	Gnews
Totalscore		-0.01	0.04	0.09	0.11	0.04	-0.12	0.04	0.21	0.19	0.03
G_experience	0.02		0.38	0.44	0.11	0.19	0.03	-0.14	-0.01	0.00	-0.02
F_experience	0.04	0.49		0.18	0.08	0.66	0.07	-0.14	-0.04	-0.03	0.03
Nfirm	0.10	0.32	0.14		0.19	0.15	0.10	-0.03	0.03	0.02	-0.03
NIndustry	0.10	0.18	0.07	0.17		0.07	-0.22	-0.04	0.00	0.03	-0.03
For_frequency	0.03	0.17	0.52	0.12	0.03		0.19	-0.04	-0.10	-0.09	0.05
Broker_size	-0.10	0.01	0.10	0.04	-0.20	0.21		-0.05	-0.05	-0.06	0.01
Accuracy	0.04	-0.15	-0.12	-0.01	-0.04	-0.03	-0.07		-0.04	-0.05	0.05
Release	0.22	0.02	-0.03	0.03	0.01	-0.11	-0.05	-0.04		0.66	0.06
Guide	0.21	0.02	-0.03	0.02	0.04	-0.10	-0.07	-0.05	0.66		0.00
Gnews	0.03	0.01	0.03	-0.03	-0.03	0.06	-0.01	0.06	0.06	0.00	

Note:

^a Pearson correlations are reported in the lower triangle; Spearman correlations are reported in the upper triangle. Correlations significant at a minimum of 10% level are in bold.

^b InText_length is the natural log of number of words in the report. Fog, Flesch, and Flesch_Kincaid are measures of readability. Totalscore is the mean value of standardized scores of Fog, Flesch, and Flesch_Kincaid across all sample reports. G_experience is the number of years for which analyst i supplied at least one forecast. F_experience is the number of years through year t for which analyst i supplied at least one forecast for firm j. NFirm is the number of companies analyst i followed in year t. NIndustry is the number of industries analyst i followed in year t. For_frequency is analyst i's forecast frequency for firm j. Broker_size is the number of analysts employed by the brokerage that employs the analyst. Accuracy is analyst i's historical forecast accuracy, measured as the standardized ranking of the analyst's average absolute forecast errors (|FE|) during 1986 to 2007, for each firm. Release is a dichotomous variable that equals one if there is an earnings release issued within five days before the analyst report date; zero otherwise. Guide is a dichotomous variable that equals one if there is an earnings guidance issued within five days before the analyst report date; zero otherwise. Gnews is a dichotomous variable that equals one if the variable Forecast > 0; zero otherwise.

Table 3. Market Reaction to Analyst Forecast ReportsThe following table reports the coefficient of the regression ^a:

$$CAR[-1, +1] = a + b*Totalscore + c*News + d*Totalscore*News + e*MVE + d*MTB + f*Momentum + e^b$$

Panel A: Dependent variable = CAR[-1, +1]

	Full sample	GNews = 1	GNews = 0
	Coef.	Coef.	Coef.
Intercept	-0.0039 (-0.39)	0.0275 * (1.79)	-0.0351 ** (-2.48)
Totalscore	0.0002 *** (3.53)	0.00004 (0.29)	0.0001 (1.24)
News	0.0862 ** (2.55)	0.0046 (0.06)	0.0435 (1.10)
Totalscore* News	0.0004 (0.68)	0.0029 ** (2.21)	0.0001 (0.08)
MVE	-0.0009 (-0.85)	-0.0012 (-0.78)	0.0019 (1.37)
MTB	-0.00001 (-0.01)	-0.0015 (-1.07)	-0.0008 (-0.55)
Momentum	0.0385 *** (6.77)	0.0059 (0.78)	0.0585 *** (7.28)
N	2,930	1,226	1,704
Adj. R ²	11.11%	9.42%	6.72%

Table 3 Continued)

Panel B. Dependent variable = CAR[-1, +1]								
	Horizon <= Median				Horizon > Median			
	Total	GNews = 1	Gnews = 0		Total	GNews = 1	GNews = 0	
	Coef.	Coef.	Coef.		Coef.	Coef.	Coef.	
Intercept	-0.0216 (-1.59)	0.0032 (0.14)	-0.0441 ** (-2.42)		0.0164 (1.14)	0.0436 ** (1.99)	-0.0109 (-0.46)	
Totalscore	0.0004 *** (3.66)	0.0001 (0.76)	0.0001 (0.66)		0.0001 (1.30)	-0.00004 (-0.21)	0.0001 (0.82)	
News	0.0653 * (1.86)	0.0080 (0.08)	0.0353 (0.88)		0.1384 (1.58)	0.0762 (0.56)	0.0752 (0.51)	
Totalscore* News	0.0004 (0.76)	0.0032 ** (2.05)	-0.0001 (-0.20)		0.0002 (0.14)	0.0017 (0.74)	0.0004 (0.17)	
MVE	-0.0002 (-0.16)	-0.0001 (-0.05)	0.0028 (1.61)		-0.0013 (-0.89)	-0.0016 (-0.76)	0.0004 (0.16)	
MTB	0.0011 (0.67)	0.0017 (0.77)	-0.0017 (-0.78)		-0.0022 (-1.55)	-0.0044 ** (-2.39)	-0.0017 (-0.79)	
Momentum	0.0531 *** (6.54)	0.0046 (0.44)	0.0797 *** (7.25)		0.0188 ** (2.43)	0.0023 (0.20)	0.0273 ** (2.46)	
N	1,525	577	948		1,405	649	756	
Adj. R ²	13.14%	9.71%	9.33%		10.14%	9.17%	4.47%	

Table 3 (Continued)**Panel C. Dependent variable = CAR[-1, +1]**

	<u>BidAskSpread > Median</u>			<u>BidAskSpread <= Median</u>		
	Total	GNews=1	GNews=0	Total	GNews=1	GNews=0
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Intercept	-0.0068 (-0.38)	0.0202 (0.72)	-0.0078 (-0.30)	0.0100 (1.22)	0.0055 (0.40)	-0.00003 (0.00)
Totalscore	0.0005 *** (3.33)	-0.00002 (-0.07)	0.0001 (0.42)	0.0001 (1.02)	0.0002 ** (2.26)	-0.000004 (-0.05)
News	0.1486 *** (2.82)	-0.0681 (-0.55)	0.1196 * (1.95)	-0.0030 (-0.11)	0.1322 ** (2.01)	-0.0278 (-0.82)
Totalscore* News	0.0002 (0.18)	0.0047 ** (2.43)	-0.0007 (-0.67)	0.0003 (0.64)	-0.0018 (-1.64)	0.0002 (0.46)
MVE	-0.0022 (-1.22)	0.0032 (1.18)	-0.0032 (-1.28)	-0.0011 (-1.29)	-0.0013 (-1.03)	-0.00002 (-0.02)
MTB	-0.0003 (-0.22)	-0.0061 *** (-3.02)	0.0002 (0.08)	0.0015 (1.26)	0.0015 (1.04)	0.0008 (0.43)
Momentum	0.0477 *** (5.82)	0.0051 (0.45)	0.0706 *** (6.36)	0.0041 (0.77)	-0.0043 (-0.60)	0.0109 (1.33)
N	1,465	575	890	1,465	651	814
Adj. R ²	17.71%	10.56%	9.80%	1.42%	2.96%	1.29%

Table 3 (Continued)**Panel D. Dependent variable = CAR[-1, +1]**

	<u>High Innovative Forecasts</u>						<u>Low Innovative Forecasts</u>					
	Total	GNews=1		GNews=0			Total	GNews=1		GNews=0		
	Coef.	Coef.		Coef.			Coef.	Coef.		Coef.		
Intercept	-0.004 (-0.31)	0.038 (2.34)	**	-0.061 (-3.45)	***		0.005 (0.24)	-0.091 (-1.83)	*	0.022 (1.00)		
Totalscore	0.000 (3.34)	***	0.000 (-0.42)		0.000 (1.60)		0.000 (1.79)	*	0.001 (2.61)	**	0.000 (-0.27)	
News	0.120 (3.49)	***	-0.002 (-0.03)		0.042 (1.05)	*	-0.143 (-2.68)	**	0.189 (0.98)	**	-0.095 (-1.74)	*
Totalscore* News	0.000 (-0.16)		0.003 (2.01)	**	0.000 (-0.18)		0.003 (2.90)	***	0.000 (-0.04)		0.001 (1.02)	
MVE	-0.002 (-1.40)		-0.002 (-1.04)		0.002 (1.39)		0.000 (0.04)		0.007 (1.62)		0.000 (-0.15)	
MTB	0.001 (0.68)		-0.001 (-0.70)	***	0.001 (0.52)		-0.002 (-1.04)		-0.009 (-2.65)	**	-0.004 (-1.43)	
Momentum	0.044 (6.59)	***	0.014 (1.67)		0.054 (5.14)	***	0.010 (0.87)		-0.085 (-3.83)	***	0.040 (3.03)	***
N	2,200		1,047		1,153		662		145		517	
Adj. R ²	14.95%		8.89%		7.96%		2.25%		23.42%		1.99%	

Note:

^a. Industry fixed effects are included in regressions with robust standard errors. T-values are reported in parentheses. ***, **, and * represent significance at 1%, 5% and 10% levels, respectively.

^b. CAR[-1, +1] is the cumulative abnormal return on day -1 to day +1, where day 0 is the analyst report date. The estimation is based on market model using returns during day -131 and day -31.

Totalscore is the mean value of standardized scores of Fog, Flesch, and Flesch_Kincaid across all sample reports. News is the difference between the analyst's EPS forecast issued on the report date (for the most recent period) and mean value of analysts' consensus forecast for the same firm and the same period, scaled by the absolute value of the consensus forecast. MVE is the logarithm of market value of equity. MTB is market value of equity plus book value of liabilities scaled by total assets. Momentum is 180-day buy-and-hold return up to a report date. GNews is a dichotomous variable that equals one if the variable News > 0; zero otherwise. Horizon is the number of days between forecast date and forecast period end date (i.e., forecast period end date – report date). BidAskSpread is the bid-ask spread of the analyst report issuing month, scaled by month-end stock price.

Table 4 Ex-post Forecast AccuracyThe following table reports the coefficient of the regression ^a:

$$|FE| = a + b * Totalscore + Controls + e.$$
^b

This is the Pre-Published Version

	FE		FE		FE	
	Full Sample		GNews=1		GNews=0	
	Coef.		Coef.		Coef.	
Intercept	4.3606	***	4.2433	***	4.3559	***
	(8.50)		(5.47)		(6.39)	
Totalscore	0.0013		-0.0056	*	0.0061	*
	(0.55)		(-1.74)		(1.92)	
MVE	-0.2836	***	-0.2982	***	-0.2870	***
	(-7.14)		(-5.47)		(-5.30)	
MTB	-0.2128	***	-0.1344	***	-0.2104	***
	(-7.50)		(-3.51)		(-5.42)	
RET_volatility	3.2810	*	-1.9339		4.7988	**
	(1.94)		(-0.83)		(2.00)	
EPS_volatility	-0.0728		0.0560		-0.1707	
	(-0.44)		(0.18)		(-0.92)	
Leverage	0.0432		0.5020	***	-0.1618	
	(0.42)		(3.16)		(-1.30)	
Egrowth	0.0245		0.2532	***	-0.0759	
	(0.45)		(2.61)		(-1.15)	
PredictE_loss	1.3353	**	-1.2383		1.9565	***
	(2.20)		(-1.33)		(2.80)	
N	2,930		1,226		1,704	
Adj. R ²	13.45%		26.91%		11.69%	

Note:

^aIndustry fixed effects are included in regressions with robust standard errors. T-values are reported in parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels, respectively.

^b|FE| is the absolute value of ex-post forecast error, where the forecast error is defined as the difference between the analyst's forecast issued on the report date and the actual realized earnings for the forecast period, scaled by stock price at the end of the fiscal quarter. Totalscore is the mean value of standardized scores of Fog, Flesch, and Flesh_Kincaid across all sample reports. GNews is a dichotomous variable that equals one if the variable News > 0; zero otherwise. MVE is the logarithm of market value of equity. MTB is market value of equity plus book value of liabilities scaled by total assets. RET_volatility is standard deviation of preceding 15-months stock returns for a specific quarter. EPS_volatility is standard deviation of EPS during the preceding 16 quarters. Leverage is the ratio of total assets to stockholders equity for the most recent quarter. Egrowth is earnings growth, measured as the change of income before extraordinary items for the most recent quarter, scaled by income before extraordinary items in the preceding quarter. PredictE_loss is a dichotomous variable that equals one if the analyst predicted earnings are negative, and zero otherwise.

Table 5. Determinants of ReadabilityThe following table reports the coefficient of the regression ^a:

This is the Pre-Published Version

$$Totalscore = a + b*Analyst\ Factors + c*Report\ Factors + d*Firm\ Factors + e.^b$$

	Expected Sign	Full sample Coef.		GNews = 1 (A) Coef.		GNews = 0 (B) Coef.	
Intercept		108.2633 *** (7.38)		123.8687 *** (4.86)		96.2497 *** (4.92)	
<i>Analyst Characteristics</i>							
G_experience	+	-0.0336 (0.35)		-0.0950 (0.65)		0.0332 (0.26)	
F_experience	+	0.5905 *** (3.35)		0.5983 ** (2.24)		0.6549 *** (2.69)	
NFirm	-	0.2977 *** (4.62)		0.3215 *** (3.29)		0.2546 *** (2.92)	
NIndustry	-	0.2055 (0.63)		0.2664 (0.51)		0.1345 (0.31)	
For_frequency	+	0.1184 * (1.75)		0.1466 (1.53)		0.0710 (0.73)	
Broker_size	+	-0.0072 (1.41)		-0.0232 *** (2.83)		0.0042 (0.63)	
Accuracy	+	0.2196 (0.14)		5.9675 ** (2.39)		-4.0573 ** (2.07)	
<i>Report Characteristics</i>							
Release	+	8.6480 *** (9.16)		8.5363 *** (5.22)		7.4102 *** (6.26)	
Guide	+	4.9286 *** (4.84)		2.9947 * (1.95)		7.1765 *** (5.27)	
REC	+/-	2.1289 *** (3.49)		1.5898 * (1.68)		2.1596 *** (2.63)	
Team	-	-2.5678 *** (3.08)		-3.0840 ** (2.26)		-2.3762 ** (2.22)	
Nreport	+/-	0.0274 (0.30)		0.1066 (0.72)		-0.0059 (0.05)	
lnText_length	-	-10.3611 *** (16.84)		-10.6419 *** (10.55)		-10.0283 *** (12.87)	
<i>Firm Characteristics</i>							
MVE	+/-	0.0754 (0.20)		-0.0477 (0.07)		-0.0282 (0.06)	
MTB	+/-	1.2269 *** (4.78)		1.0893 *** (2.71)		1.1945 *** (3.45)	
RET_volatility	-	8.8625 (0.71)		11.1167 (0.56)		4.0890 (0.25)	
EPS_volatility	-	1.5025 (1.45)		-1.3415 (0.77)		3.1922 ** (2.41)	
MAQR	-	-2.7509 *** (3.33)		-3.3322 ** (2.55)		-1.6945 (1.56)	
SEOQR	-	7.2091 (0.95)		-3.9959 (0.91)		35.5662 *** (13.13)	
NBSEG	-	-1.0314 (1.11)		-1.9389 (1.30)		-0.3545 (0.29)	
NGSEG	-	-1.5702 * (1.68)		-2.8080 * (1.81)		-0.8271 (0.69)	
NITEM	-	-0.0702 (0.02)		-1.4844 (0.29)		1.5593 (0.40)	
SI	-	74.9828 (0.92)		81.9494 (0.59)		94.5241 (0.92)	

AGE	+	0.0777 **	0.0267	0.1206 **
		(2.25)	(0.53)	(2.55)
DLW	-	-2.2077 ***	-3.6063 ***	-1.9956 **
		-(2.66)	-(2.57)	-(1.33)
N		2,930	1,226	1,704
Adj. R ²		26.32%	27.32%	26.00%

Note:

- a. Industry fixed effects are included in the regressions with robust errors. T-values are reported in parentheses. ***, ** and * represent significance at 1%, 5% and 10% levels, respectively.
- b. G_experience is the number of years for which analyst i supplied at least one forecast during year t. F_experience is the number of years through year t for which analyst i supplied at least one forecast during year t for firm j. NFirm is the number of companies analyst i followed in year t. NIndustry is the number of industries analyst i followed in year t. For_frequency is analyst i's forecast frequency for firm j. Broker_size is the number of analysts in brokerage that employs the analyst. Accuracy is analyst i's historical forecast accuracy, measured as the standardized ranking of the analyst's average absolute forecast errors (|FE|) during 1986 and 2007 for each firm; |FE| is scaled by absolute value of actual earnings. Release is a dichotomous variable that equals one if there is an earnings release issued within five days before the analyst report date; zero otherwise. Guide is a dichotomous variable that equals one if there is an earnings guidance issued within five days before the analyst report date; zero otherwise. REC is the difference between the analyst's recommendation issued on the report date and mean value of analysts' consensus recommendation for the same firm, scaled by the consensus recommendation and multiplied by -1. Recommendation codes follow those used in I/B/E/S (i.e., 1 = strong buy; 2 = buy; 3 = hold; 4 = underperform; 5 = sell.) News is the difference between the analyst's EPS forecast issued on the report date (for the most recent period) and mean value of analysts' consensus forecast for the same firm and the same period, scaled by the absolute value of the consensus forecast. GNews is a dichotomous variable that equals one if the variable News > 0; zero otherwise. Team is a dichotomous variable that equals one if a specific report lists more than one analyst, and zero otherwise. NReport is the number of analyst reports issued on the same day for the same firm (including those whose analysts cannot be identified in I/B/E/S.) lnText_length is the natural log of number of words in the report. Totalscore is the mean value of standardized scores of Fog, Flesch, and Flesh_Kincaid across all sample reports. MVE is the logarithm of market value of equity. MTB is market value of equity plus book value of liability scaled by total assets. RET_volatility is the standard deviation of 15-month stock returns for a specific quarter. EPS_volatility is the standard deviation of EPS during the preceding 16 quarters. MAQR is a dichotomous variable that equals one if the firm appears in the SDC M&A database as an acquirer in the quarter; zero otherwise. SEOQR is a dichotomous variable that equals one if the firm appears in the SDC database for issuing common equity in the quarter; zero otherwise. NBSEG is the number of business segments. NGSEG is the number of geographic segments. NITEM is the logarithm of the number of non-missing items in Compustat. SI is special items scaled by total assets. Age is firm age since the firm first appeared in CRSP monthly stock file. DLW is a dichotomous variable that equals one if a firm was incorporated at Delaware, and zero otherwise.