The Information Content of Management Earnings Forecasts: An Analysis of Hard versus Soft Information*

Paul Brockman**
Department of Finance
Robert J. Trulaske, Sr. College of Business
University of Missouri
Phone: (573) 884-1562
E-mail: brockmanp@missouri.edu

Jim Cicon
Department of Finance
Robert J. Trulaske, Sr. College of Business
University of Missouri
Phone: (573) 882-4505
E-mail: jecdhf@mizzou.edu

*The authors would like to thank Financial Economics Seminar participants at the University of Missouri for their valuable comments.
**Corresponding author.

The Information Content of Management Earnings Forecasts: An Analysis of Hard versus Soft Information

Abstract

We examine the announcement effects of hard (quantitative) and soft (qualitative) information contained in management earnings forecasts. Consistent with previous studies, we confirm a positive relation between the earnings surprise component of the announcement and the magnitude of the abnormal return. In contrast with previous studies, we examine the impact of soft information on abnormal returns while controlling for the impact of hard information. We use textual analysis to identify and measure the level of two specific soft information variables (Net Optimism and Certainty) present in each earnings forecast. Our empirical results show that soft information plays a significant role in explaining the magnitude of abnormal returns. In fact, soft information generally explains more of the announcement effect of earnings forecasts than hard information. Overall, our findings contribute to the growing literature that examines the economic significance of qualitative information.

JEL classifications: G30; G31; G34
Key words: Management earnings forecasts; Soft information; Hard information; Textual analysis; Voluntary disclosures

The Information Content of Management Earnings Forecasts: An Analysis of Hard versus Soft Information

1. Introduction

Although many previous studies have examined the corporate announcement effects of hard or quantitative information on firm values, few studies to date have analyzed the announcement effects of soft or qualitative information. Whether the corporate announcement is related to earnings, dividends, security issuances, or restructuring, previous studies have relied almost exclusively on changes in numerical values to explain the abnormal returns associated with the announcement. However, most corporate announcements contain far more qualitative terms than quantitative terms. We show that qualitative information comprises roughly 99% of the terms in a typical management forecast. Does this preponderance of qualitative terms simply reflect "filler material" that is used to plug the gaps between hard numbers? Or do these qualitative terms provide value-relevant information for current and potential investors? In this study, we attempt to shed light on these important questions by examining the announcement effects of both hard and soft information contained in management earnings forecasts. Our empirical results show that investors rely on the information content of soft data at least as much as they rely on the information content of hard data.

Our study focuses on management earnings forecasts for two main reasons. First, previous studies show that management forecasts significantly affect stock prices (Pownall, Wasley, and Waymire (1993)), stock market liquidity (Coller and Yohn (1997)), and analyst forecast revisions (Baginski and Hassell (1990)).[1] Since forecasted earnings announcements clearly contain value-relevant information, we can examine the announcement effects attributable to soft information versus hard information. More specifically, we will measure the

[1] See Hirst, Koonce, and Venkataraman (2006) for a review of the management earnings forecast literature.

incremental effect of soft information on announcement abnormal returns after controlling for the effect of hard information. Second, management earnings forecasts represent voluntary disclosures. Recent research (e.g., Aboody and Kasznik (2000), Cheng and Lo (2006), Brockman, Khurana, and Martin (2008)) has shown that managers exercise considerable leeway in the timing and content of their voluntary disclosures. The form and format of such disclosures are less formulaic than their required disclosure counterparts. Managers can freely choose the hard-soft information mix in their announcements. This attribute allows us to investigate whether managers use these additional degrees of freedom to convey value-relevant information through qualitative terms.

Several recent studies examine the relation between the company's reported earnings and soft information (e.g., Li (2006), Engelberg (2007), Davis, Piger, and Sedor (2007), Tetlock, Saar-Tsechansky, Macskassy (2008)). Among these studies, our paper is most closely related to that of Demers and Vega (2008).[2] These authors examine the announcement effects of soft versus hard information using over 20,000 earnings announcements. They find that soft information generates a significant impact on stock prices during earnings announcements above and beyond the impact of hard information. Interestingly, they also find that it takes longer for investors to fully interpret the soft information content in earnings announcements than the hard information content. It is apparently easier to read and interpret a numerical section than a narrative section.

In this study, we examine the market's interpretation of hard and soft information contained in voluntary disclosures. We collect over 15,000 management earnings forecasts from 1994 to 2007 to examine the role that soft information plays in the pricing of forecast announcements. We estimate three-day cumulative abnormal returns (CARs) using standard

[2] Other studies that examine qualitative aspects of information announcements include Das and Chen (2007), Tetlock (2007), and Das, Martinez-Jerez, and Tufano (2005).

event study methodology and use these CARs as our dependent variable. We then fit a series of regression models to extract the explanatory power of hard versus soft information contained in management forecast announcements.

Similar to previous studies analyzing actual earnings releases (e.g., Demers and Vega (2008)), we regress CARs on both hard and soft information variables. We use a modified version of standardized unexpected earnings (i.e., standardized unexpected forecasted earnings, or SUFEs) as our hard information variable, and Net Optimism and Certainty as our soft information variables.[3] Net Optimism, as defined by Diction (a standard textual analysis program), includes language that is positively associated with praise, satisfaction, or inspiration; and negatively associated with blame, hardship, or denial. Certainty includes language that is positively associated with tenacity, leveling, collectives, or numerical terms; and negatively associated with ambivalence, self-reference, or variety.

Our regression results show that the relation between three-day announcement CARs and hard information (SUFE) is positive and significant, consistent with previous findings and expectations. More importantly, our regression results confirm that Net Optimism and Certainty have positive and significant effects on three-day announcement CARs – over and above any hard information effects. Moreover, the estimated coefficients on our standardized independent variables show that our soft information variables, Net Optimism and Certainty, generally have a stronger impact on announcement returns than our hard information variable, SUFE. To the best of our knowledge, these findings represent the first empirical evidence that the market prices the qualitative information contained in management earnings forecasts.

In subsequent tests, we examine whether the relation between announcement returns and soft information is linear, concave, or convex. We find some evidence of concavity for one of

[3] More precisely, our definition of SUE is the standardized unexpected earnings *forecast*; that is, the difference between the management forecast and the actual earnings figure one year earlier. For annual forecasts, we compare the manager's forecast with actual annual earnings one year earlier. For quarterly forecasts, we compare the manager's forecast with actual quarterly earnings one year earlier.

the soft information variables, Certainty. While there is a diminishing marginal impact of Certainty on announcement returns, we find no such nonlinearity for Net Optimism. Next, we examine the possibility of interaction effects between hard information and soft information. It is not clear on an ex ante basis whether hard and soft information behave as compliments or substitutes to one another. It is also possible that hard and soft information have no significant interaction effects. Our empirical results reveal evidence in favor of a substitution effect between SUFE and Certainty, and between SUFE and Net Optimism. These findings suggest that managers can use positive soft information to offset some of the announcement effects of negative hard information.

We also extend our analysis and examine the announcement impact of hard and soft information as a function of the announcing firm's information environment (as proxied by firm Size). As hypothesized, we find that the interaction between Size and Net Optimism is negative and significant. We find the same negative and significant relation for the interaction between Size and Certainty. These findings suggest that soft information plays a particularly important role for small firms with less abundant information environments. Next, we examine the interaction effects between our information variables and alternative proxies for the firm's information environment (Analyst Coverage and Numerical Terms), and between our information variables and firm-specific characteristics (Intangibility and M/B Ratio). We find weak evidence of negative interactions between our information variables and the alternative proxies for the firm's information environment; consistent evidence of negative interactions between the information variables and Intangibility; and no significant interactions between the information variables and growth opportunities.

Lastly, we perform robustness tests using alternative proxies for qualitative information. Loughran and McDonald (2009) point out that commonly-used English word classification lists might be inappropriate for finance-related research. Words such as "liability" or "depreciation"

have more negative connotations in common usage than in finance-specific usage. To investigate whether our findings are sensitive to word-list selection, we apply their alternative word lists to construct five qualitative variables, including "positive" (Fin-Pos) and "uncertainty" (Fin-Unc). Our regression results confirm a positive and significant relation between earnings announcement returns and Fin-Pos, and a negative and significant relation between earnings announcement returns and Fin-Unc. These findings suggest that our empirical results are robust to the use of alternative word lists.

Our study makes two primary contributions. First, we add to the growing literature showing that corporate announcements contain significant, value-relevant soft as well as hard information. Until recently, most previous research focused exclusively on the value relevance of hard information. Our results support the idea that soft information variables should be included as independent variables whenever a researcher attempts to identify the relevant explanatory variables related to corporate announcement CARs. Second, our results provide a better understanding of the role played by management earnings forecasts. This type of corporate announcement is particularly interesting because, unlike earnings or dividend announcements, management forecast announcements are completely voluntary. Managers not only decide on when to announce their forecasts, they also determine the hard-versus-soft information mix in these announcements. Such a setting provides a unique opportunity to analyze the use and value relevance of soft information.

The rest of this study proceeds as follows. We describe our data and methods of analysis in Section 2. This section also discusses our testable hypotheses. We present and analyze our empirical findings in Section 3, and then provide a brief conclusion in Section 4.

2. Data and methods of analysis

2.1. Data description

Our study analyzes the soft information found in the text of company-issued-guidance press releases. We use Lexis Nexis academic as the source of these press releases. Lexis Nexis organizes press releases into many different databases; we limited our search to the 'Company Press Releases' database. However, this database has many news providers and press releases are often duplicated among these different news providers. Therefore, we further limit our search to the following news providers: PR Newswire (US), Business Wire, and Canada Newswire.

Once we identify the database, and the news providers within the database, we continue by searching for candidate announcements over the period of 1994 to 2007. Our search criteria analyze the 'subject' line and the 'headline' of each potential candidate. Using the subject line, we keep only those press releases which fall into the following categories: ("earnings preannouncements" or "earnings projections or forecasts") and not ("interim financial results"). We exclude the 'interim financial results' hits because press releases in this category tend to originate from non-company sources (e.g., analyst commentary). We then scan the headlines of the remaining documents and retain only those that contain the following terms: (guidance or forecast or predict or prediction or projection or forward or future or expect or expected or expectation or update or anticipate or reaffirm) and ((cash w/s flow) or earn or earning or share or stock or EPS or fiscal or price or revenue or results or profit or profitability or targets or growth or raise or lower or upward or upwardly or downward or downwardly). The 'and' connector between the two groups of words means that one of the first group of words AND one of the second group of words must appear in the headline in order for the press release to be retained.

The results of the Lexis-Nexis search yields 74,603 candidate press releases. However, not all of these press releases are relevant. For example, some files are duplicates and other files are commentary by third parties subsequent to the original press release. In order to ensure

relevancy, we match the candidate press releases with the First Call 'Company Issued Guidelines' database. Since the 74,603 press releases downloaded from Lexis-Nexis are in PDF format, they are not readily identifiable by a machine. Therefore, we wrote custom software to scan and extract the ticker symbol and the announcement date from each press release. We match this ticker symbol and date to records in the First Call database. We retain only those press releases which match the First Call ticker symbol and which occur on the same day as the First Call guidance release. Where more than one press release occurs on the same day for the same company, we keep the earliest. We also discard press releases that do not contain at least 100 words.

After this step we are left with 15,010 press releases. We randomly sample these press releases to ensure they were relevant matches. We next run custom algorithms to 'clean up' the press releases. We start this clean-up by first removing all punctuation from the press releases. Hyphenated words are concatenated. Numerical terms are counted and then removed, with a record of their count retained. If tables are found, they are removed – although we keep record of their presence with a dummy variable. Words which do not occur more than two times in a given document are discarded. As mentioned previously, each document has to have at least 100 words in it to be included in our sample. We also truncate press releases which are longer than 750 words. This effectively removes company-description and safe-harbor paragraphs in the longer press releases. Shorter press releases do not typically contain these boiler plate paragraphs.

The data sample is then divided into three groups. The first group contains only quarterly announcements and totals 5,242 documents. The second group contains only annual announcements and totals 9,769 documents. The third group aggregates both the quarterly and annual groups and totals 15,011 documents.

2.2. Methods of analysis and hypothesis development

Our methodology analyzes each press release for the presence of semantic features, specifically, for the presence of Net Optimism and Certainty. We base our analysis on Diction, a language processing software package which has been used extensively in prior research to analyze political and economic speeches, corporate reports, and earnings announcements. Diction analyzes a document by counting the frequency of specific key words. It then compares these counts to norms calculated from a broad sample of English text. This sample of text includes 20,000 documents consisting of speeches, newspaper editorials, business reports, scientific documents, telephone conversations, etc..

Diction defines Net Optimism as language endorsing some person, group, concept or event or highlighting their positive accomplishments (Carroll (2000)). The program generates an aggregate standardized score for Net Optimism by counting words in defined categories, and then combines the standardized scores for each category as shown in the following equation:

$$NetOpt = (praise + satisfaction + inspiration) - (blame + hardship + denial) \quad (1)$$

Diction defines Certainty as language indicating resoluteness, inflexibility, completeness and a tendency to speak ex cathedra. The program generates an aggregate standardized score for Certainty by counting words in defined categories, and then combines the standardized scores for each category as shown in the following equation:[4]

$$Certainty = (tenacity + leveling + collectives + numerical\_terms) - (ambivalence + self\_reference + variety) \quad (2)$$

These two semantic features serve as the soft information variables in our study.

Our analysis proceeds as follows. We begin by testing whether soft information has explanatory power with respect to cumulative abnormal returns. We test this hypothesis by running the following regression:

[4] In addition to standardizing our qualitative information variables using Diction's out-of-sample mean and variance values, we replicate all empirical tests using in-sample mean and variance values. Our empirical results and conclusions are not sensitive to the standardization procedure.

$$\sum_{i=-1}^{+1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \varepsilon_{jt} \quad (3)$$

This expression contains both hard and soft explanatory variables. SUFE is the standardized unexpected forecasted earnings, the hard or numerical information variable. We construct SUFE by first subtracting the management earnings forecast of firm *j* for time period *t* (e.g., third quarter of 2002) from the firm's actual (reported) earnings in time period *t-1* (e.g., third quarter of 2001). We use this difference as the unexpected component in the manager's forecast (i.e., the unexpected forecasted earnings (UFE)). We then calculate standardized unexpected forecasted earnings (SUFE) by subtracting the sample mean and dividing by the sample standard deviation. Net Optimism and Certainty are the soft information explanatory variables.[5] We use their standardized value in all regressions as well. We expect the coefficients on all three treatment variables to be positive and statistically significant.

We revise regression model (3) by adding squared terms for our soft, qualitative variables. This approach allows us to analyze the possibility of nonlinear relations between Net Optimism (Certainty) and abnormal announcement returns.[6] We examine these possibilities using the following regression:

$$\sum_{i=-1}^{1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 NetOpt_{j,t}^2 + \beta_4 Certainty_{j,t} + \beta_5 Certainty_{j,t}^2 + \varepsilon_{j,t} \quad (4)$$

Negative (positive) and significant coefficients on the squared qualitative terms would be evidence of a concave (convex) relation between soft information and announcement effects. To the extent that there exists any nonlinearities in these relations, concavity appears to be more likely than convexity. That is, we expect that the marginal impact of Net Optimism or Certainty

[5] See Appendix A for more detailed descriptions of all dependent and independent variables.
[6] Although our main interest is with squared qualitative (Net Optimism and Certainty) terms, we also estimate a model with a squared quantitative (SUFE) term.

on announcement returns will be a constant or diminishing (but not increasing) relation over higher levels of Net Optimism or Certainty.

We again alter our benchmark regression model (3) in order to directly examine the relation between hard information and soft information. It is possible that soft information compliments hard information (i.e., a positive interaction effect), has no relation to hard information (i.e., no interaction effect), or acts as a substitute to hard information (i.e., a negative interaction effect). Since we do not have strong priors regarding an expected relation and the existing literature offers little to no guidance, we treat this analysis as an open-ended empirical issue. We estimate regression model (5) as follows:

$$\sum_{i=-1}^{1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \beta_4 SUFE_{j,t} NetOpt_{j,t} + \beta_5 SUFE_{j,t} Certainty_{j,t} + \varepsilon_{j,t} \quad (5)$$

Positive (negative) and significant coefficients on the interaction terms would be evidence that hard and soft information behave as compliments (substitutes).

Next, we include size as an explanatory, along with interaction terms, to determine if the impact of soft information on abnormal earnings is sensitive to the information environment of the firm as proxied by Size (market value of equity). Following Demers and Vega (2008), we estimate the following regression:

$$\sum_{i=-1}^{+1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \beta_4 Size_{j,t} + \beta_5 SUFE_{j,t} Size_{j,t} + \beta_6 NetOpt_{j,t} Size_{j,t} + \beta_7 Certainty_{j,t} Size_{j,t} + \varepsilon_{j,t} \quad (6)$$

We expect that SUFE, Net Optimism, and Certainty will have less impact on abnormal returns as the information environment improves (i.e., as firm size increases). We therefore expect to find negative and significant coefficients on all three interaction terms.[7]

Finally, we revise regression model (6) by including alternative proxies for the information environment of the firm (Analyst Coverage and Numerical Terms), as well as fundamental firm characteristics (Intangibility and M/B Ratio). Regression model (7) can be expressed as follows:

$$\sum_{i=-1}^{+1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \beta_4 X_{j,t} + \beta_5 SUFE_{j,t} X_{j,t} + \beta_6 NetOpt_{j,t} X_{j,t} + \beta_7 Certainty_{j,t} X_{j,t} + \varepsilon_{j,t} \quad (7)$$

where $X_{j,t}$ represents (in succession) Analyst Coverage, Numerical Terms, Intangibility, and the M/B Ratio. Analyst Coverage is defined as the logarithm of the number of analysts that post an earnings estimate for the current period, as reported by I/B/E/S; Numerical Terms is the standardized value of the number of numerical terms in the announcement document divided by the total number of words in the document; Intangibility is calculated as intangible assets divided by total assets; and the B/M Ratio is the company's market value of equity divided by its book value of equity.[8]

We expect both Analyst Coverage and Numerical Terms to exhibit negative interaction effects with our hard and soft information variables. That is, the more analysts covering a firm or the more numerical information contained in a disclosure (i.e., the better the information environment), the less the expected price impact from SUFE, Net Optimism, or Certainty. Similarly, we expect a negative interaction effect between Intangibility and our information

[7] In addition, we include industry control variables based on one digit SIC codes in both regression models (6) and (7).

[8] Following Demers and Vega (2008), we also examine the level of R&D expenditures (as a percent of total assets) and a high tech industry dummy variable as interaction terms in regression model (7).

variables. All else equal, the more intangible the firm's assets, the less the expected price impact from SUFE, Net Optimism, or Certainty. Stated differently, firms with more tangible assets will have larger CARs for a given level of SUFE, Net Optimism, or Certainty. In addition to assets in place, we also examine growth opportunities as proxied by the M/B Ratio. We do not make specific predictions for the interactions between growth opportunities and our information variables. We therefore treat this part of the analysis as an open empirical question.

3. Empirical results

3.1. Summary statistics

We report descriptive statistics for our main variables of interest in Table 1. These variables include Optimism, Pessimism, Net Optimism, Certainty, SUFE, Numerical Terms, CARs, Size, Analyst Coverage, Intangibility, and the M/B ratio. Panel A's sample statistics include the combined earnings forecasts, Panel B's sample statistics are based on quarterly earnings forecasts only, and Panel C's sample statistics are based on annual earnings forecasts only. In each panel, we report the variable's mean, median, standard deviation, 25th percentile, and 75th percentile.

The mean (median) Optimism value in Panel A is -0.9760 (-1.2648). These negative values show that management earnings forecasts tend to be less optimistic than their respective benchmark (i.e., a cross-section of English texts). The mean (median) Pessimism value is -1.4674 (-1.6743), showing that management forecasts also tend to be less pessimistic than their benchmark.[9] These results are consistent with the notion that managers are concerned about litigation risk (Hirst, Koonce, and Venkataraman (2006)) and therefore tend to use conservative language in their voluntary disclosures. When we combine these two measures, we find that Net

[9] Optimism increases with the use of terms related to praise, satisfaction, and inspiration, while Pessimism increases with the use of terms related to blame, hardship, and denial (see Section 2.1. and Appendix A).

Optimism has a mean (median) value of 0.4914 (0.3752). Since Net Optimism is one of our main explanatory (soft) variables, it is also important to note that it displays considerable variation with a value of -1.1068 at the 25th percentile and 1.9520 at the 75th percentile. Our other main (soft) variable is Certainty, which has a mean (median) value of 16.4979 (15.8046).[10] This explanatory variable also displays considerable variation with a value of 12.2737 at the 25th percentile and 19.9572 at the 75th percentile.

Forecasted earnings surprises, or SUFEs, represent our main hard variable of interest. In Panel A, its mean (median) value is 0.0593 (0.0413), and its 25th (75th) percentile value is -0.0900 (0.1850). If the information content of management forecasts consists only of hard (i.e., numerical) data, then SUFE will have substantial explanatory power over abnormal announcement returns – and our soft variables, Net Optimism and Certainty, will have little to no explanatory power. We include summary statistics for Size since previous studies use firm size as a proxy for asymmetric information, and interact Size with both hard and soft explanatory variables (see Demers and Vega (2008)). The mean firm Size in our sample is $1.1 billion ($e^{13.9026}$ x 1000), and the median Size is $994 million ($e^{13.81}$ x 1000).

Next, we report summary statistics for our dependent variable (3-day cumulative abnormal returns (CARs)), as well as for our main control variables (Numerical Terms, Analyst Coverage, Intangibility, and the M/B ratio). Mean and median CARs are negative (-0.0146 and -0.0027, respectively), consistent with previous studies (e.g., Cheng and Lo (2006), Brockman, Khurana, and Martin (2008)). The Numerical Terms' mean (median) of 0.0127 (0.0092) illustrates that roughly 1% of the entire forecast announcement is dedicated to the provision of numerical values (i.e., hard data). In other words, soft data constitute 99% of the forecast announcement's content. Our sample firms have a mean (median) of 2.6558 (2.000) financial

[10] Certainty increases with the use of terms related to resoluteness, inflexibility, and completeness (see Section 2.1. and Appendix A).

analysts following their companies, and the mean (median) level of Intangibility is 0.1288 (0.0370).

We find similar results for our sample of quarterly forecasts in Panel B and annual forecasts in Panel C, with some notable exceptions. The Net Optimism mean and median values for annual forecasts are considerable larger than their quarterly forecast counterparts. The Net Optimism annual forecast mean (median) is 0.7501 (0.5843) in Panel C, compared to the quarterly forecast mean (median) of 0.0092 (-0.0876) in Panel B. Managers appear to be more optimistic in annual forecasts than in quarterly forecasts based on their hard information as well. The SUFE mean (median) for annual forecasts is 0.0832 (0.0625) in Panel C, compared to 0.0148 (0.0000) for quarterly forecasts in Panel B. While managers tend to be less optimistic when providing quarterly forecasts, they also tend to be less certain about quarterly forecasts. The mean (median) value for Certainty is only 14.8567 (14.1322) for quarterly forecasts, compared to 17.3784 (16.8953) for annual forecasts.

3.2. Correlations

In Table 2, we report correlations among our dependent and control variables. We again include separate panels for the combined, quarterly, and annual forecasts for comparison purposes. Correlations among most of our independent variables are less than 10%, suggesting that multicollinearity is not a major concern. We therefore focus our discussion on those variables with correlations greater than 10%. Net Optimism and Size have positive correlations of 19%, 8%, and 22% in Panels A, B, and C, respectively. Certainty and Numerical Terms have positive correlations of 19%, 38%, and 18% in Panels A, B, and C, and Certainty and CARs have positive correlations of 11%, 9%, and 9%, respectively. Although somewhat weaker, SUFE is also positively correlated with CARs (8%, 11%, and 5%). Analyst Coverage is negatively correlated with CARs (-9%, -7%, and -11%), and positively correlated with Size (42%. 48%, and

41%) and Intangibility (14%, 12%, and 10%). Finally, we note that Intangibility and Numerical Terms are negatively correlated (-15%, -16%, and -8%).

3.3. Multivariate analysis: Benchmark case

In this section, we estimate regression model (3) and report the results in Table 3. We are interested in examining the impact of hard and soft information content in management forecasts on announcement returns. We hypothesize that both hard and soft information play a significant role in impounding value-relevant news into stock market prices; that is, we expect the estimated coefficients for each of the three explanatory variables to be positive and significant. Because our explanatory variables have been standardized, we can also compare the relative impact of each variable on abnormal announcement returns.

The empirical findings in Table 3 support our hypothesis that the market prices qualitative (i.e., soft) as well as quantitative (i.e., hard) information. In Panel A, we show that the estimated coefficient for our hard information variable, SUFE, is positive and significant (coefficient = 0.00908; *t*-value = 9.76), consistent with expectations that the larger the forecasted earnings surprise, the larger the abnormal return. More importantly, we find that the estimated coefficients for our soft information variables, Net Optimism and Certainty, are positive and significant (coefficients = 0.00946 and 0.01179; *t*-values = 10.16 and 12.67, respectively). These results confirm that the market prices the soft information in management earnings forecasts, particularly as this information relates to Net Optimism and Certainty. We also note that the estimated coefficient magnitudes and related *t*-statistics of the soft variables are larger than those of the hard variable. Certainty's coefficient, for example, is roughly 30% larger than SUFE's coefficient. To the best of our knowledge, these results represent the first empirical evidence in the literature that verify the importance of soft information in management forecasts.

In Panels B and C of Table 3, we repeat our analysis and examine quarterly and annual forecasts separately. In both panels, the estimated coefficients for SUFE are positive and

significant. More importantly, the estimated coefficients for Net Optimism and Certainty are consistently positive and significant. While the estimated coefficient for SUFE is larger than those of the soft variables in Panel B using quarterly forecast data, the estimated coefficients for Net Optimism and Certainty are both larger than the estimated coefficient for SUFE in Panel C using annual forecast data. Overall, the results in Table 3 show that not only does the market price soft information, but that it generally puts more weight on the soft information contained in announcements than on the hard, numerical information contained in these same announcements.

3.4. Analysis of nonlinearities

In Table 4, we report the results from regression model (4) that includes squared soft information terms. We established in the previous section that both Net Optimism and Certainty have positive and significant relations with the abnormal returns surrounding management forecasts. In this section, we further examine these relations by introducing squared terms into regression model (3) for each of the soft information variables.[11] A positive (negative) and significant coefficient on a squared term suggests that the relation between abnormal returns and our soft information variable is convex (concave). The interpretation of a convex (concave) relation is that the positive and significant relation documented in Table 3 increases (decreases) at higher levels of the soft information variable.

Across all three panels in Table 4, we find that the estimated coefficients for both the hard (SUFE) and soft (Net Optimism and Certainty) information variables remain positive and significant after including the squared terms. More importantly, we find that the estimated coefficient for the squared Certainty term is negative and significant in all three panels. These findings show that there is a diminishing marginal effect of Certainty on abnormal returns around management earnings forecasts (i.e., a concave relation). In contrast, we find no significant

[11] In untabulated results, we also examine nonlinearities in the relation between our hard information variable, SUFE, and CARs. We find that the estimated coefficient for SUFE squared is negative and significant, similar to but weaker than the estimated coefficient for Certainty squared. However, including a SUFE squared term into regression model (4) does not alter our conclusions with respect to Net Optimism squared or Certainty squared.

relation between squared Net Optimism and abnormal returns – suggesting that changes in Net Optimism have similar effects on abnormal returns regardless of the starting level of Net Optimism. So, while Certainty experiences diminishing returns, Net Optimism displays constant returns.

3.5. Hard and soft information: Compliments or substitutes?

In this section, we estimate regression model (5) to examine the relation between hard information and soft information. We are interested in determining whether hard and soft information behave as compliments or substitutes. It is also possible that these two information types have no significant interaction effects. We report our empirical results in Table 5.

We find a significant substitution effect for both Net Optimism and Certainty using the combined data in Panel A. The estimated coefficient for the interaction between Net Optimism and SUFE is -0.00151 (*t*-value = 1.94), suggesting that Net Optimism is more effective when SUFE is relatively small. We find similar, yet stronger, results for the interaction between Certainty and SUFE, with an estimated coefficient of -0.00373 (*t*-value = 3.94). Again, we interpret this result as evidence of a substitution effect between hard information and soft information. Certainty has a stronger impact on the market's reaction to earnings forecasts when SUFE is relatively small. When we separate our combined data into quarterly forecasts (Panel B) and annual forecasts (Panel C), we find that the substitution effect observed in Panel A is only significant for the annual forecasts.

Overall, our Table 5 results show that hard information and soft information exhibit substitution effects, particularly at the level of annual earnings forecasts. In contrast, we find no evidence of a complimentary relation for quarterly, annual, or combined earnings forecasts.

3.6. Interactions with firm size

In Table 6, we add firm Size to our baseline regression (3) and examine its interaction with both hard and soft information variables (i.e. regression model (6)). Since Size proxies for

the information environment of the firm, we expect to find that the firm Size mitigates the impact of hard and soft information on abnormal returns (i.e., negative coefficients on interaction terms). The empirical results in Table 6 consistently support this hypothesis for both soft and hard information variables in Panels A, B, and C.[12]

More specifically, we show in Panel A that the estimated coefficient for the interaction between Size and SUFE is -0.00672 (*t*-value = -6.79). Turning to our soft information variables, we find that the estimated coefficient for the interaction between Size and Net Optimism is -0.00370 (*t*-value = -4.33), and the estimated coefficient for the interaction between Size and Certainty is -0.00759 (*t*-value = -8.10). We note firm Size has the greatest effect on Certainty, precisely what one would expect if Size is a good proxy for the information environment of the firm. These results add further evidence to our central claim that soft information contains significant explanatory power with respect to managerial announcement effects. The negative interaction-term coefficients show that the impact of both hard and soft information depends on the firm's information environment. The influence of soft information is similar to, yet generally stronger than, the impact of hard information.

Similar findings are presented in Panel B for quarterly data and Panel C for annual data. In Panel B, that the estimated coefficient for the interaction between Size and SUFE is -0.00563 (*t*-value = -2.92), the estimated coefficient for the interaction between Size and Net Optimism is -0.00477 (*t*-value = -2.76), and the estimated coefficient for the interaction between Size and Certainty is -0.00986 (*t*-value = -4.64). Similarly, the results in Panel C show that the interactions between Size and all information variables are negative and statistically significant. Overall, we interpret these findings as evidence that soft information is particularly important when the firm's information environment is relatively opaque.

3.7. Analysis of additional interaction effects

[12] All empirical results in Table 6 include one digit SIC control (dummy) variables.

In Table 7, we estimate regression model (7) by examining the interactions between our main variables of interest (SUFE, Net Optimism, and Certainty) and alternative proxies for the firm's information environment (Analyst Coverage and Numerical Terms) – as well as for proxies of fundamental firm characteristics (Intangibility and the M/B Ratio). Our main objective is to examine whether the price impact of Net Optimism and Certainty on earnings forecast announcements is influenced by these specific-firm attributes.[13]

As discussed above in our hypothesis development section, we expect negative interaction effects between both Analyst Coverage and Numerical Terms and our information variables. That is, the better the firm's information environment (i.e., the more the Analyst Coverage and Numerical Terms), the less the price impact that we expect from SUFE, Net Optimism, or Certainty. Although most of the estimated coefficients for these interaction terms are negative in Panel A (Analyst Coverage), none of them is statistically significant. And in Panel B (Numerical Terms), we find only weak evidence of negative interaction effects. The estimated coefficient for the interaction between Numerical Terms and Certainty is -0.00187 (*t*-value = -2.14) using the combined forecasts, and -0.00455 (*t*-value = -3.04) using the quarterly forecast. None of the remaining interaction terms is significant. These results suggest that Size (as used in Table 6) is a better proxy for the firm's information environment than either Analyst Coverage or Numerical Terms. Alternatively, one can interpret these results as showing that the impact of SUFE and Net Optimism on announcement CARs is simply not sensitive to Analyst Coverage or Numerical Terms.

In Panels C and D, we examine the interaction between firm characteristics and SUFE, Net Optimism, and Certainty. The results in Panel C confirm our expectation of negative interactions between Intangibility and these information variables. All else equal, firms with more tangible assets exhibit a greater price impact from SUFE, Net Optimism, and Certainty on

[13] As in Table 6, all empirical results in Table 7 include one digit SIC control (dummy) variables.

their announcement CARs. The estimated coefficient for the interaction between Intangibility and SUFE is -0.00317 (*t*-value = -3.12) based on the combined forecasts; it is -0.00246 (*t*-value = -2.68) for Net Optimism, and -0.00446 (*t*-value = -4.89) for Certainty. Most of these interaction-term coefficients are also negative and significant using quarterly and annual forecasts. In Panel D, we find no significant interactions between the M/B Ratio and our information variables. Taken together, our Panel C and D results suggest that the impact of SUFE, Net Optimism, and Certainty on a firm's announcement returns is sensitive to its assets in place, but insensitive to its future growth opportunities.[14]

3.8. Robustness tests: Alternative definitions

In a recent paper, Loughran and McDonald (2009) point out that commonly-used word classification lists, designed to identify words with positive or negative connotations in public texts, might be inappropriate for finance-related research. While words such as liability or depreciation have negative implications in common usage, they have neutral implications in financial usage. Loughran and McDonald propose five alternative word lists based on word counts (and meanings) contained in 10-K filings. These word lists include (1) positive (Fin-Pos), (2) uncertain (Fin-Unc), (3) litigious (Fin-Lit), (4) strong modal words (MW-Strong), and (5) weak modal words (MW-Weak). The first two qualitative variables, Fin-Pos and Fin-Unc, correspond closely to our qualitative variables, Net Optimism and Certainty, respectively. We therefore use Fin-Pos and Fin-Unc as alternative proxies for our qualitative information variables.[15] We expect to find a positive (negative) and significant relation between announcement returns and Fin-Pos (Fin-Unc).

[14] Our main findings and conclusions reported in Table 7 are unaltered by the inclusion of Size as an additional explanatory variable in regression model (7).

[15] We also estimate regressions using all five qualitative terms (Fin-Pos, Fin-Unc, Fin-Lit, MW-Strong, and MW-Weak) plus SUFE as our independent variables. The signs, magnitudes, and significance levels for the Fin-Pos and Fin-Unc coefficients from these regressions are very similar to those reported in Table 8.

Our findings in Table 8 based on alternative soft information proxies (i.e., Fin-Pos and Fin-Unc) confirm our earlier Table 3 results based on Net Optimism and Certainty. In Panel A, we find that the estimated coefficient for the hard information variable, SUFE, is positive and significant (coefficient = 0.00949; *t*-value = 10.13), consistent with expectations and our Table 3 results. More importantly, we show that the estimated coefficient for Fin-Pos is positive and statistically significant (coefficient = 0.00587; *t*-value = 6.20), and the estimated coefficient for Fin-Unc is negative and statistically significant (coefficient = -0.00536; *t*-value = 5.70).

In Panels B and C of Table 3, we repeat our analysis and examine quarterly and annual forecasts separately. In both panels, the estimated coefficients for SUFE are positive and significant. The estimated coefficients for Fin-Pos in Panels B and C are also both positive and significant. Fin-Unc has a negative and significant coefficient in Panel C using annual data, but its coefficient is negative and insignificant in Panel B using quarterly data. Overall, the results in Panels B and C are consistent with expectations and similar to those in Panel A.

More generally, our Table 8 results show that the market prices the soft information in management earnings forecasts, consistent with our previous findings using Net Optimism and Certainty. Announcement returns are significantly larger when the qualitative text of the announcement reveals managerial optimism and certainty – even after controlling for the text's quantitative information. Taken together, both our main results and these additional robustness checks support the conclusion that qualitative information is value relevant.

4. Conclusions

In this study, we examine the announcement effects of hard (quantitative) and soft (qualitative) information contained in management earnings forecasts. Consistent with previous studies, we confirm a positive relation between the earnings surprise component (i.e., SUFE) of the announcement and the magnitude of the abnormal return. In contrast with previous studies,

we examine the impact of soft information on abnormal returns while controlling for the impact of hard information.

More specifically, we collect over 15,000 management earnings forecasts from 1994 to 2007 to examine the role that soft information plays in the pricing of forecast announcements. We estimate three-day cumulative abnormal returns (CARs) using standard event study methodology. In our benchmark case, we regress these CARs on both hard and soft information variables. Our empirical results confirm that the relation between CARs and hard information (SUFE) is positive and significant, consistent with expectations. More importantly, our results confirm that Net Optimism and Certainty have positive and significant effects on these same CARs – an incremental effect over and above any hard information effects. To the best of our knowledge, these are the first empirical results showing that the qualitative data in management forecasts contain value-relevant information. Without this finding, it is difficult to explain why rational managers dedicate roughly 99% of the content of their voluntary disclosures to qualitative information.

In subsequent empirical tests, we find that Certainty exhibits a concave relation to the announcement impact of management forecasts. While there is a diminishing marginal impact of Certainty on announcement returns, we find no such nonlinearity for Net Optimism. We also examine the relation between hard information and soft information. It is possible that hard and soft information act as compliments or substitutes to one another. Our results reveal consistent evidence in favor of a substitution effect between SUFE and Certainty, as well as between SUFE and Net Optimism based on our full sample. Separating our sample into quarterly and annual forecasts shows that the substitution effect is generated from annual management forecasts.

Next, we extend this analysis and examine the announcement impact of hard and soft information as a function of the announcing firm's information environment (as proxied by firm Size). As hypothesized, we find that the interaction between Size and Net Optimism is negative

and significant for quarterly, annual, and combined forecast announcements. We find the same negative and significant relation for the interaction between Size and Certainty. These findings show that soft information plays an especially important role for small firms with less transparent information environments.

We then examine the interaction effects between our information variables and other proxies for the firm's information environment (Analyst Coverage and Numerical Terms), and between our information variables (SUFE, Net Optimism, and Certainty) and firm-specific characteristics (Intangibility and M/B Ratio). We find weak evidence of negative interaction effects between our information variables and the alternative proxies for the firm's information environment. In contrast, we find consistent evidence of negative and significant interaction effects between our information variables and Intangibility, and no significant interaction effects between our information variables and the firm's growth opportunities.

Lastly, we perform robustness tests using alternative proxies for qualitative information. We apply alternative word lists designed by Loughran and McDonald (2009) and construct their qualitative terms, positive (Fin-Pos) and uncertainty (Fin-Unc). Our subsequent regression results confirm a positive (negative) and significant relation between management earnings forecast announcements and Fin-Pos (Fin-Unc). These findings suggest that our empirical results are robust to alternative word list definitions.

Our study makes two primary contributions to the finance literature. First, we add to the growing literature showing that corporate announcements contain value-relevant soft, as well as hard, information. Second, our results provide a better understanding of the role played by management earnings forecasts. Such announcements are particularly interesting because, unlike earnings or dividend announcements, management earnings forecasts are voluntary. Managers decide what to announce (i.e., hard versus soft information) and when to announce it. This setting provides a unique opportunity to analyze the value relevance of soft information on

market valuations. Overall, our study provides new evidence on the relative importance of hard versus soft information in the context of voluntary disclosures.

References

Aboody, D., Kasznik, R., 2000. CEO stock option awards and the timing of voluntary corporate disclosures. *Journal of Accounting and Economics* 29, 73-100.

Baginski, S., Hassell, J., 1990. The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revision. *The Accounting Review* 65, 175-190.

Brockman, P., Khurana, I., Martin, X., 2008. Voluntary disclosures around share repurchases, *Journal of Financial Economics* 89, 175-191.

Cheng, Q., Lo, K., 2006. Insider trading and voluntary disclosures. *Journal of Accounting Research* 44, 815-848.

Coller, M., Yohn, T., 1997. Management forecasts and information asymmetry: An examination of bid-ask spreads. *Journal of Accounting Research* 35, 181-191.

Das, S., Chen, M., 2007. Yahoo! For Amazon: Sentiment extraction from small talk on the web. *Management Science* 53, 1375-1388.

Das, S., Martinez-Jerez, A., Tufano, P., 2005. E-information: A clinical study of investor discussion and sentiment. *Financial Management* 34, 103-137.

Davis, A., Piger, J., Sedor, L., 2007. Beyond the numbers: Managers' use of optimistic and pessimistic tone in earnings press releases. Working paper, University of Washington.

Demers, E., Vega, C., 2008. Soft information in earnings announcements: News or noise? Working paper, INSEAD.

Engelberg, J., 2007. Costly information processing: Evidence from earnings announcements. Working paper, Northwestern University.

Hirst, E., Koonce, L., Venkataraman, S., 2006. Management earnings forecasts: A review and framework. Working paper, University of Texas.

Kasznik, R., Lev, B., 1995. To warn or not to warn—management disclosures in the face of an earnings surprise. *The Accounting Review* 70, 113–34.

Li, Feng, 2006. Do stock market investors understand the risk sentiment of corporate annual reports? Working paper, University of Michigan.

Loughran, T., McDonald, B., 2009, When is a liability not a liability?, Working paper, University of Notre Dame.

Pownall, G., Wasley, C., Waymire, G., 1993. The stock price effects of alternative types of management earnings forecasts. *The Accounting Review* 68, 896-912.

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

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

Appendix A
Summary of Variable Definitions

Variable	Definition
Optimism	Language which is optimism increasing (praise + satisfaction + inspiration). This value is normalized using means and standard deviations computed from a cross-section of English text.
Pessimism	Language which is optimism decreasing (blame + hardship + denial). This value is normalized using means and standard deviations computed from a cross-section of English text.
net optimism (not standardized)	Optimism minus pessimism
NetOpt (standardized)	$NetOpt_{j,t} = \frac{netoptimism_{j,t} - \mu_{netoptimism}}{\sigma_{netoptimim}}$ $netoptimism_{j,t}$ is the net optimism for firm 'j' in quarter 't', $\mu_{netoptimism}$ is the mean net optimism across all firms across all quarters, and $\sigma_{netoptimism}$ is the standard deviation in net optimism across all firms across all quarters
certainty (not standardized)	Language indicating resoluteness, inflexibility, and completeness (tenacity + leveling + collectives + numerical_terms) – (ambivalence + self_reference + variety). Note that we follow Demers and Vega (2008) add the numerical_terms rather than subtract them. This is to accommodate the different meaning of numerical terms in financial text. This value is normalized using means and standard deviations computed from a cross-section of English text.
Certainty (standardized)	$Certainty_{j,t} = \frac{certainty_{j,t} - \mu_{certainty}}{\sigma_{certainty}}$ $certainty_{j,t}$ is the certainty for firm 'j' in quarter 't', $\mu_{certainty}$ is the mean certainty across all firms across all quarters, and $\sigma_{certainty}$ is the standard deviation in certainty across all firms across all quarters
UFE	Unexpected Forecasted Earnings: $UFE_{j,t} = CIG_{j,t} - EPS_{j,t-4}$ $CIG_{j,t}$ is the new earnings forecast for firm 'j' in quarter 't', and $EPS_{j,t-4}$ is the actual earnings for firm 'j' in last years quarter 't-4'.
SUFE	Standardized Unexpected Forecasted Earnings: $SUFE_{j,t} = \frac{UFE_{j,t} - \mu_{UFE}}{\sigma_{UFE}}$ $UFE_{j,t}$ is defined above μ_{UFE} is the mean of UFE across all firms across all quarters σ_{UFE} is the standard deviation of UFE across all firms across all quarters

Summary of Variable Definitions (continued)

Variable	Definition
CARs	We use EVENTUS to calculate four different CARs based on combinations of two market indices (CRSP Equally Weighted & CRSP Value Weighted) and two benchmarks (Market-Adjusted & Market Model). Our estimation period parameters are as follows: • EST (end before event date): -30 • MINESTN (minimum estimation length): 3 • ESTLEN (maximum estimation length): 100 Our event windows are defined as follows: • Event window 1: PRE = -1, POST = 1 • Event window 2: PRE = 2, Post = 65 We only report results for the 'Market' model using the 'CRSP Equal Weighted' index. In untabulated tests we run all four combinations of model/index and find no substantive difference in results. We find that event window matters. In untabulated tests we find that our results are not significant in the longer event window. We believe that this is because the longer event window extends past the end of the quarter, after the actual earnings are released. Once the investors have actual earnings information, the old soft information becomes irrelevant to them.
Size	The natural logarithm of the firm's market capitalization defined as shares outstanding ('shrout') times share price ('prc). This data is pulled from CRSP.
Analyst	The natural logarithm of the number of analysts that posted an earnings estimate for the current quarter, as reported on I/B/E/S.
Intangibility	Intangibility is defined as the ratio of Intangible assets ('intanq') divided by total assets ('atq').
Numerical	The normalized value of the number of numerical terms in the document divided by the total number of words in the document.
M/B	Market value divided by book value. We calculate it as the ratio of ('shrout' * 'prc)/ceqq. 'ceqq' is defined as 'common/ordinary equity – total' and is used in accounting literature as a measure of book value.
Industry	Firms grouped by SIC codes.
HiTech	Hi Tech firms are not limited to one particular industry. Therefore we compute Hi Tech as follows: hitech equal to 1 if SIC is 3570-3579, 3622, 3660-3692, 3694-3699, 3810-3839, 7370-7372, 7373-7379, 7391, 8730-8734
R&D	Research and Development Ratio: We use COMPUSTAT data to estimate the R&D ratio as R&D expenses ('xrdy') over total expenditures ('xsgay' plus 'xrdy').

Table 1: Descriptive Statistics

In this table we present summary statistics for the following variables. *Optimism* which consists of words in the following categories: praise, satisfaction and inspiration. *Pessimism* which consists of words in the following categories: blame, hardship and denial. *NetOpt* which is Optimism less Pessimism. *Certainty* which consists of words indicating resoluteness, inflexibility and completeness. *CARs* which are computed over 3 days around the announcement date. *SUFE* which is the standardized unexpected earnings forecast. *Analyst* which is the number of analysts covering firm j in quarter t. *Intangibility* which is intangible assets divided by total assets. *M/B* which is market value divided by book value. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. The sample was divided three ways. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	average	median	std dev	25 percentile	75 percentile
Optimism	-0.9760	-1.2648	1.9051	-2.2613	-0.0303
Pessimism	-1.4674	-1.6743	1.5277	-2.5027	-0.6453
NetOpt	0.4914	0.3752	2.5976	-1.1068	1.9520
Certainty	16.4979	15.8046	6.0432	12.2737	19.9572
SUFE	0.0593	0.0413	0.4882	-0.0900	0.1850
CARs	-0.0146	-0.0027	0.1151	-0.0539	0.0403
Size	13.9026	13.8100	1.8046	12.7210	15.0832
Analyst	2.6558	2.0000	2.9375	1.0000	4.0000
Intangibility	0.1288	0.0370	0.1777	0.0000	0.2086
Numerical	0.0127	0.0092	0.0127	0.0056	0.0150
M/B	4.1385	2.1840	172.1769	1.4228	3.4966

Panel B: Quarterly Data

	average	median	std dev	25 percentile	75 percentile
Optimism	-1.1408	-1.4007	1.8865	-2.4238	-0.1780
Pessimism	-1.1500	-1.3514	1.6580	-2.2738	-0.2582
NetOpt	0.0092	-0.0876	2.6638	-1.6208	1.5206
Certainty	14.8567	14.1322	5.2965	11.3063	17.5216
SUFE	0.0148	0.0000	0.4981	-0.1700	0.1400
CARs	-0.0334	-0.0160	0.1406	-0.0909	0.0410
Size	13.4189	13.3261	1.8337	12.1414	14.5087
Analyst	2.7646	2.0000	3.2035	0.0000	4.0000
Intangibility	0.0841	0.0000	0.1546	0.0000	0.0983
Numerical	0.0166	0.0124	0.0148	0.0078	0.0202
M/B	3.0772	2.0061	14.7103	1.2688	3.2364

Table 1 (continued): Descriptive Statistics

Panel C: Annual Data

	average	median	std dev	25 percentile	75 percentile
Optimism	-0.8876	-1.2004	1.9092	-2.1625	0.0547
Pessimism	-1.6377	-1.8251	1.4242	-2.5985	-0.8860
NetOpt	0.7501	0.5843	2.5239	-0.8006	2.1659
Certainty	17.3784	16.8953	6.2327	13.0358	20.9277
SUFE	0.0832	0.0625	0.4811	-0.0500	0.2050
CARs	-0.0045	0.0013	0.0972	-0.0377	0.0401
Size	14.1621	14.0703	1.7341	12.9971	15.3051
Analyst	2.5974	2.0000	2.7827	1.0000	4.0000
Intangibility	0.1528	0.0742	0.1846	0.0000	0.2586
Numerical	0.0106	0.0079	0.0108	0.0050	0.0124
M/B	4.7080	2.2974	213.1535	1.5126	3.6035

Table 2: Correlation Matrix

In this table we present summary statistics for the following variables. *Optimism* which consists of words in the following categories: praise, satisfaction and inspiration. *Pessimism* which consists of words in the following categories: blame, hardship and denial. *NetOpt* which is Optimism less Pessimism. *Certainty* which consists of words indicating resoluteness, inflexibility and completeness. *CARs* which are computed over 3 days around the announcement date. *SUFE* which is the standardized unexpected earnings forecast. *Analyst* which is the number of analysts covering firm j in quarter t. *Intangibility* which is intangible assets divided by total assets. *M/B* which is market value divided by book value. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. The sample was divided three ways. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	NetOpt	Certainty	SUFE	CARs	Size	Analyst	Intangibility	Numerical	M/B
NetOpt	1								
Certainty	0.01	1							
SUFE	0.02	0.03	1						
CARs	0.09	0.11	0.08	1					
Size	0.19	0.06	0.03	0.09	1				
Analyst	0.05	-0.04	0.01	-0.09	0.42	1			
Intangibility	0.09	0.07	0.05	0.10	0.14	0.02	1		
Numerical	-0.06	0.19	-0.03	-0.01	-0.07	0.01	-0.15	1	
M/B	0.00	0.00	0.01	0.00	0.01	0.00	-0.01	0.01	1

Panel B: Quarterly Data

	NetOpt	Certainty	SUFE	CARs	Size	Analyst	Intangibility	Numerical	M/B
NetOpt	1								
Certainty	-0.01	1							
SUFE	0.04	0.03	1						
CARs	0.08	0.09	0.11	1					
Size	0.08	0.07	0.02	0.11	1				
Analyst	0.02	0.02	0.01	-0.07	0.48	1			
Intangibility	0.05	0.03	0.07	0.10	0.12	0.03	1		
Numerical	-0.01	0.38	0.01	0.05	-0.01	-0.01	-0.16	1	
M/B	0.02	0.02	0.02	0.01	0.06	0.00	0.02	0.00	1

Panel C: Annual Data

	NetOpt	Certainty	SUFE	CARs	Size	Analyst	Intangibility	Numerical	M/B
NetOpt	1								
Certainty	-0.01	1							
SUFE	0.00	0.01	1						
CARs	0.06	0.09	0.05	1					
Size	0.22	-0.01	0.02	0.03	1				
Analyst	0.08	-0.06	0.01	-0.11	0.41	1			
Intangibility	0.07	0.03	0.03	0.07	0.10	0.02	1		
Numerical	-0.05	0.18	-0.03	-0.01	-0.04	0.01	-0.08	1	
M/B	-0.01	0.00	0.01	0.00	0.00	0.01	-0.01	0.02	1

Table 3: The Effect of Hard and Soft Information on Disclosure Period CARs
In this table we estimate the following equation:

$$\sum_{i=-1}^{1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \varepsilon_{j,t}$$

Where $SUFE_{j,q}$ is the standardized unexpected forecasted earnings for firm 'j' in quarter 'q'. $NetOpt_{j,t}$ is the standardized net optimism for firm 'j' in quarter 't'. $Certainty_{j,t}$ is the standardized certainty for firm 'j' in quarter 't'. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. The sample was divided three ways. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	coefficient	t-stat	p-value			
SUFE	0.00908	9.76	0.0000	***	Adj R2:	2.44%
NetOpt	0.00946	10.16	0.0000	***	f-stat:	125.30
Certainty	0.01179	12.67	0.0000	***		
Intercept	-0.01461	-15.71	0.0000	***		

Panel B: Quarterly Data

	coefficient	t-stat	p-value			
SUFE	0.01406	7.45	0.0000	***	Adj R2:	2.46%
NetOpt	0.01079	5.75	0.0000	***	f-stat:	44.82
Certainty	0.01386	6.31	0.0000	***		
Intercept	-0.02639	-12.88	0.0000	***		

Panel C: Annual Data

	coefficient	t-stat	p-value			
SUFE	0.00520	5.23	0.0000	***	Adj R2:	1.40%
NetOpt	0.00638	6.33	0.0000	***	f-stat:	47.15
Certainty	0.00819	8.62	0.0000	***		
Intercept	-0.00659	-6.62	0.0000	***		

Table 4: The effect of Squared Soft Information on Disclosure Period CARs
In this table we estimate the following equation:

$$\sum_{i=-1}^{1} AR_{j,t+1} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 NetOpt_{j,t}^2 + \beta_4 Certainty_{j,t} + \beta_5 Certainty_{j,t}^2 + \varepsilon_{j,t}$$

Where $SUFE_{j,t}$ is the standardized unexpected forecasted earnings for firm 'j' in quarter 't'. $NetOpt_{j,t}$ is the standardized net optimism for firm 'j' in quarter 't'. $Certainty_{j,t}$ is the standardized certainty for firm 'j' in quarter 't'. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. The sample was divided three ways. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	coefficient	t-stat	p-value			
SUFE	0.00905	9.72	0.0000	***	Adj R2:	2.58%
NetOpt	0.00936	9.89	0.0000	***	f-stat:	80.13
NetOpt^2	-0.00007	-0.13	0.8940			
Certainty	0.01410	13.49	0.0000	***		
Certainty^2	-0.00250	-4.91	0.0000	***		
Intercept	-0.01204	-10.35	0.0000	***		

Panel B: Quarterly Data

	coefficient	t-stat	p-value			
SUFE	0.01416	7.50	0.0000	***	Adj R2:	2.55%
NetOpt	0.01088	5.79	0.0000	***	f-stat:	28.34
NetOpt^2	0.00156	1.57	0.1167			
Certainty	0.01516	6.69	0.0000	***		
Certainty^2	-0.00291	-2.20	0.0276	*		
Intercept	-0.02526	-9.70	0.0000	***		

Panel C: Annual Data

	coefficient	t-stat	p-value			
SUFE	0.00517	5.21	0.0000	***	Adj R2:	1.54%
NetOpt	0.00651	6.11	0.0000	***	f-stat:	31.44
NetOpt^2	-0.00035	-0.63	0.5285			
Certainty	0.01034	9.32	0.0000	***		
Certainty^2	-0.00193	-3.86	0.0001	***		
Intercept	-0.00449	-3.75	0.0002	***		

Table 5: Interaction effects between Hard and Soft Information

In this table we estimate the following regression:

$$\sum_{i=-1}^{1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \beta_4 SUFE_{j,t} NetOpt_{j,t} + \beta_5 SUFE_{j,t} Certainty_{j,t} + \varepsilon_{j,t}$$

Where $SUFE_{j,t}$ is the standardized unexpected forecasted earnings for firm 'j' in quarter 't'. $NetOpt_{j,t}$ is the standardized net optimism for firm 'j' in quarter 't'. $Certainty_{j,t}$ is the standardized certainty for firm 'j' in quarter 't'. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. The sample was divided three ways. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	coefficient	t-stat	p-val			
SUFE	0.00874	9.36	0.0000	***	Adj R2:	2.55%
NetOpt	0.00940	10.09	0.0000	**	f-stat:	79.14
Certainty	0.01187	12.76	0.0000	***		
SUFE*NetOpt	-0.00152	-1.94	0.0527	.		
SUFE*Certainty	-0.00373	-3.94	0.0001	***		
Intercept	-0.01445	-15.54	0.0000	***		

Panel B: Quarterly Data

	coefficient	t-stat	p-val			
SUFE	0.01357	6.48	0.0000	***	Adj R2:	2.45%
NetOpt	0.01097	5.80	0.0000	***	f-stat:	27.21
Certainty	0.01371	6.24	0.0000	***		
SUFE*NetOpt	0.00107	0.65	0.5160			
SUFE*Certainty	-0.00240	-1.13	0.2590			
Intercept	-0.02638	-12.88	0.0000	***		

Panel C: Annual Data

	coefficient	t-stat	p-val			
SUFE	0.00570	5.65	0.0000	***	Adj R2:	1.50%
NetOpt	0.00644	6.39	0.0000	***	f-stat:	30.61
Certainty	0.00837	8.80	0.0000	***		
SUFE*NetOpt	-0.00163	-1.90	0.0574	.		
SUFE*Certainty	-0.00295	-2.97	0.0030	**		
Intercept	-0.00658	-6.61	0.0000	***		

Table 6: The Effect of Market Capitalization and Hard and Soft Information on Disclosure Period CARs

We interact size and control for industry in the baseline regression:

$$\sum_{-1=i}^{1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 SUFE_{j,t}\ Size_{j,t} + \beta_3 sNetOpt_{j,t} + \beta_4 sNetOpt_{j,t}\ Size_{j,t}$$
$$+ \beta_5 sCertainty_{j,t} + \beta_6 Certainty_{j,t}\ Size_{j,t} + \beta_7 Size_{j,t} + \varepsilon_{j,t}$$

Where $Size_{j,t}$ is the standardized natural log of the firm's market capitalization. $SUFE_{j,t}$ is the standardized unexpected forecasted earnings for firm 'j' in quarter 't'. $NetOpt_{j,t}$ is the standardized net optimism for firm 'j' in quarter 't'. $Certainty_{j,t}$ is the standardized certainty for firm 'j' in quarter 't'. We control for industry by using one digit SIC codes. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	coefficient	t-stat	p-value			
SUFE	0.00969	10.35	0.0000	***	Adj R2:	3.85%
NetOpt	0.00873	9.16	0.0000	***	f-stat:	38.41
Certainty	0.01194	12.63	0.0000	***		
SUFE	0.00623	6.31	0.0000	***		
size*SUFE	-0.00672	-6.79	0.0000	***		
size*NetOpt	-0.00370	-4.33	0.0000	***		
size*Certainty	-0.00759	-8.10	0.0000	***		
Intercept	-0.01308	-2.56	0.0105	*		

Panel B: Quarterly Data

	coefficient	t-stat	p-value			
SUFE	0.01313	6.90	0.0000	***	Adj R2:	4.51%
NetOpt	0.01009	5.28	0.0000	***	f-stat:	16.40
Certainty	0.01038	4.68	0.0000	***		
SUFE	0.00746	3.59	0.0003	***		
size*SUFE	-0.00563	-2.92	0.0035	**		
size*NetOpt	-0.00477	-2.76	0.0058	**		
size*Certainty	-0.00986	-4.64	0.0000	***		
Intercept	-0.04070	-3.64	0.0003	***		

Panel C: Annual Data

	coefficient	t-stat	p-value			
SUFE	0.00690	6.51	0.0000	***	Adj R2:	1.90%
NetOpt	0.00643	6.00	0.0000	***	f-stat:	12.79
Certainty	0.00937	9.39	0.0000	***		
SUFE	0.00240	2.19	0.0286	*		
size*SUFE	-0.00552	-4.78	0.0000	***		
size*NetOpt	-0.00188	-1.94	0.0521	.		
size*Certainty	-0.00488	-4.90	0.0000	***		
Intercept	0.00118	0.22	0.8236			

Table 7: Interaction Effects with Information Environment and Firm Characteristics

We interact firm characteristics and control for size and industry to estimate the following equation:

$$\sum_{i=-1}^{+1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 NetOpt_{j,t} + \beta_3 Certainty_{j,t} + \beta_4 X_{j,t} + \beta_5 SUFE_{j,t} X_{j,t} + \beta_6 NetOpt_{j,t} X_{j,t} + \beta_7 Certainty_{j,t} X_{j,t} + \varepsilon_{j,t}$$

Where $X_{j,t}$ is the standardized firm characteristic. $SUFE_{j,t}$ is the standardized unexpected forecasted earnings for firm 'j' in quarter 't'. $sNetOpt_{j,t}$ is the standardized net optimism for firm 'j' in quarter 't'. $Certainty_{j,t}$ is the standardized certainty for firm 'j' in quarter 't'. We control for industry by using one digit SIC codes. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

	Combined Forecasts				**Quarterly Forecasts**				**Annual Forecasts**			
	coefficient	t-stat	p-val		coefficient	t-stat	p-val		coefficient	t-stat	p-val	
					Panel A: Analyst Coverage							
SUFE	0.00932	9.98	0.000	***	0.01491	7.86	0.000	***	0.00539	5.41	0.000	***
NetOpt	0.01051	11.20	0.000	***	0.01269	6.73	0.000	***	0.00726	7.12	0.000	***
Certainty	0.01119	11.95	0.000	***	0.01275	5.79	0.000	***	0.00783	8.18	0.000	***
Analyst	-0.00908	-9.74	0.000	***	-0.01092	-5.65	0.000	***	-0.00844	-8.21	0.000	***
Analyst * SUFE	-0.00094	-1.05	0.294		-0.00144	-0.75	0.454		-0.00086	-0.92	0.357	
Analyst * NetOpt	-0.00032	-0.35	0.725		-0.00228	-1.30	0.195		0.00059	0.57	0.567	
Analyst * Certainty	0.00041	0.43	0.666		0.00030	0.15	0.883		0.00055	0.55	0.583	
Intercept	-0.01730	-3.38	0.001	***	-0.05060	-4.55	0.000	***	0.00023	0.04	0.965	
					Panel B: Numerical Terms							
SUFE	0.00909	9.75	0.000	***	0.01497	7.70	0.000	***	0.00532	5.07	0.000	***
NetOpt	0.00974	10.32	0.000	***	0.01238	6.18	0.000	***	0.00637	6.24	0.000	***
Certainty	0.01200	12.54	0.000	***	0.01350	5.60	0.000	***	0.00843	8.58	0.000	***
Numerical	-0.00149	-1.31	0.190		0.00345	1.69	0.091	.	-0.00193	-1.33	0.184	
Numerical * SUFE	0.00046	0.57	0.572		-0.00111	-0.91	0.365		0.00046	0.34	0.736	
Numerical * NetOpt	0.00009	0.11	0.916		0.00025	0.17	0.864		-0.00061	-0.56	0.578	
Numerical * Certainty	-0.00187	-2.14	0.032	*	-0.00455	-3.04	0.002	**	-0.00074	-0.68	0.496	
Intercept	-0.01591	-3.10	0.002	**	-0.04886	-4.38	0.000	***	0.00097	0.18	0.855	

Table 7: Interaction Effects with Information Environment and Firm Characteristics (Continued)

	Combined Forecasts				Quarterly Forecasts				Annual Forecasts			
	coefficient	t-stat	p-val		coefficient	t-stat	p-val		coefficient	t-stat	p-val	
					Panel C: Intangibility							
SUFE	0.00792	8.41	0.000	***	0.01281	6.26	0.000	***	0.00488	4.89	0.000	***
NetOpt	0.00911	9.76	0.000	***	0.01050	5.38	0.000	***	0.00626	6.13	0.000	***
Certainty	0.01082	11.53	0.000	***	0.01096	4.84	0.000	***	0.00839	8.64	0.000	***
Intangibility	0.01172	11.99	0.000	***	0.01539	6.64	0.000	***	0.00695	6.75	0.000	***
Intangibility * SUFE	-0.00317	-3.12	0.002	**	-0.00209	-0.87	0.385		-0.00209	-2.03	0.043	*
Intangibility * NetOpt	-0.00246	-2.68	0.007	**	-0.00555	-2.72	0.007	**	-0.00004	-0.04	0.967	
Intangibility * Certainty	-0.00446	-4.89	0.000	***	-0.00572	-2.25	0.025	*	-0.00303	-3.37	0.001	***
Intercept	-0.01664	-3.25	0.001	**	-0.04252	-3.83	0.000	***	-0.00139	-0.26	0.792	
					Panel D: M/B ratio							
SUFE	0.00918	9.84	0.000	***	0.01448	7.59	0.000	***	0.00529	5.30	0.000	***
NetOpt	0.00987	10.53	0.000	***	0.01220	6.44	0.000	***	0.00652	6.42	0.000	***
Certainty	0.01146	12.20	0.000	***	0.01287	5.65	0.000	***	0.00814	8.48	0.000	***
M/B	0.00000	0.11	0.910		-0.00014	-0.27	0.784		0.00000	0.08	0.934	
M/B * SUFE	-0.00005	-0.56	0.575		0.00022	0.43	0.671		-0.00007	-0.89	0.375	
M/B * NetOpt	-0.00001	-0.83	0.405		0.00008	0.40	0.691		-0.00001	-0.92	0.358	
M/B * Certainty	0.00000	-0.06	0.953		-0.00009	-0.14	0.892		0.00000	-0.08	0.940	
Intercept	-0.01622	-3.16	0.002	**	-0.04929	-4.42	0.000	***	0.00138	0.26	0.795	

Table 8: Loughran and McDonald (2009) definitions

In this table we estimate the following equation:

$$\sum_{i=-1}^{1} AR_{j,t+i} = \beta_0 + \beta_1 SUFE_{j,t} + \beta_2 Fin-Pos_{j,t} + \beta_3 Fin-Unc_{j,t} + \varepsilon_{j,t}$$

Where $SUFE_{j,q}$ is the standardized unexpected forecasted earnings for firm 'j' in quarter 'q'. $Fin\text{-}Pos_{j,t}$ is the standardized Loughran and McDonald (2009) positive score for firm 'j' in quarter 't'. $Fin\text{-}Unc_{j,t}$ is the standardized Loughran and McDonald (2009) uncertainty for firm 'j' in quarter 't'. The sample period spans 1994 to 2007 and includes those firms which appeared in the First Call company issued guidelines database and for which a matching company press release was found in Lexus-Nexus. The sample was divided three ways. In Panel A we report results for our entire sample of 14,938 company issued guidance press releases. In Panel B we report results for the 5,216 quarterly releases. In Panel C we report results for the 9,722 annual releases.

Panel A: Combined Data

	coefficient	t-stat	p-value			
SUFE	0.00949	10.13	0.0000	***	Adj R2:	2.44%
Fin-Pos	0.00587	6.20	0.0000	***	f-stat:	125.30
Fin-Unc	-0.00536	-5.70	0.0000	***		
Intercept	-0.01460	-15.60	0.0000	***		

Panel B: Quarterly Data

	coefficient	t-stat	p-value			
SUFE	0.01460	7.70	0.0000	***	Adj R2:	2.46%
Fin-Pos	0.00740	3.97	0.0001	***	f-stat:	44.82
Fin-Unc	-0.00211	-1.13	0.2580			
Intercept	-0.03182	-16.28	0.0000	***		

Panel C: Annual Data

	coefficient	t-stat	p-value			
SUFE	0.00519	5.21	0.0000	***	Adj R2:	1.40%
Fin-Pos	0.00501	4.90	0.0000	***	f-stat:	47.15
Fin-Unc	-0.00552	-5.42	0.0000	***		
Intercept	-0.00513	-5.21	0.0000	***		