Journal of Banking & Finance 36 (2012) 992–1011

Contents lists available at SciVerse ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Earnings conference calls and stock returns: The incremental informativeness of textual tone

S. McKay Price [a,*], James S. Doran [b], David R. Peterson [b], Barbara A. Bliss [b]

[a] Lehigh University, Perella Department of Finance, Bethlehem, PA 18015, USA
[b] Florida State University, College of Business, Tallahassee, FL 32306, USA

ARTICLE INFO

Article history:
Received 8 December 2010
Accepted 13 October 2011
Available online 19 October 2011

JEL classification:
D80
D82
G10
G12
G14
G30

Keywords:
Conference calls
Disclosure
Content analysis
Textual analysis
Stock returns
Post-earnings-announcement drift

ABSTRACT

Quarterly earnings conference calls are becoming a more pervasive tool for corporate disclosure. However, the extent to which the market embeds information contained in the tone (i.e. sentiment) of conference call wording is unknown. Using computer aided content analysis, we examine the incremental informativeness of quarterly earnings conference calls and the corresponding market reaction. We find that conference call linguistic tone is a significant predictor of abnormal returns and trading volume. Furthermore, conference call tone dominates earnings surprises over the 60 trading days following the call. The question and answer portion of the call has incremental explanatory power for the post-earnings-announcement drift and this significance is primarily concentrated in firms that do not pay dividends, illustrating differences in investor behavior based on the level of cash flow uncertainty. Additionally, we find that a context specific linguistic dictionary is more powerful than a more widely used general dictionary (Harvard IV-4 Psychosocial).

1. Introduction

Recent literature argues that earnings conference calls have become an increasingly important medium through which firms convey value relevant information to the market (Frankel et al., 1999, 2010; Kimbrough, 2005). Such disclosures help mitigate potential information asymmetry between managers and investors. Conference calls typically begin with prepared statements by management, which usually reiterate the press release, and are then opened up to questions from analysts. The dialogue between management and analysts relative to firm performance is a potentially rich information source. However, Healy and Palepu (2001) note that one of the major limitations in studying conference call voluntary disclosure, is the difficulty in measuring the extent of the disclosure.

Core (2001) contends that since call participants use natural language for this type of communication, we may be able to advance our understanding of the impact of call content by using natural language processing techniques from fields such as computer science, linguistics, and artificial intelligence. Core asserts that it is worthwhile to investigate the potential for machine-coding conference call disclosure. For example, established processes for analyzing linguistic content could be used to create proxies for the "tone", or sentiment, of disclosure (Lang and Lundholm, 2000).

Research dealing with the content analysis of disclosures is a broad field of study including, but not limited to, areas such as reading ease, determining themes, visual and structural effects, performance comparisons, choice of earnings number, and attribution of organizational outcomes (Merkl-Davies and Brennan, 2007).[1] Within content analysis, extracting the linguistic tone of a narrative using word count strategies falls within the more general realm of ascertaining meaning or themes. However, isolating tone moves beyond the analysis of *what* is being said to *how* it is being said (Pennebaker et al., 2003). Extant literature suggests that these nuanced features

* Corresponding author. Tel.: +1 610 758 4787; fax: +1 610 882 9415.
E-mail addresses: smp210@lehigh.edu (S. McKay Price), jsdoran@cob.fsu.edu (J.S. Doran), d.peterson@cob.fsu.edu (D.R. Peterson), bab04e@fsu.edu (B.A. Bliss).

[1] See Merkl-Davies and Brennan (2007) for a detailed review of content analysis studies dealing with each of the various aspects of disclosure in corporate narratives.

0378-4266/$ - see front matter
doi:10.1016/j.jbankfin.2011.10.013

of communication are responsible for the primary conveyance of meaning (Chung and Pennebaker, 2007; Davis et al., 2008).

Other channels through which meaning may be transmitted in a spoken setting include the use of both voice and body by those involved in the communication (Mehrabian, 1971, 1972). Mehrabian suggests that these additional channels encompass voice inflection, speech volume, speech rate, speech errors, facial expressions, hand gestures, posture, and movements. As an example related to the present study, concurrent work by Mayew and Venkatachalam (forthcoming) examine managerial tone of voice using a unique set of conference call audio files. Using specialized layered voice analysis software, they measure the positive and negative emotional states of managers and relate these vocal cues to stock returns. To borrow a content analysis metaphor from Hart (2001) and Pennebaker et al. (2003), this avenue of inquiry (examining tone of voice) differs from ours (examining tone of wording) like two people trying to understand a city where one drives on the streets and the other views it from a helicopter. Both obtain valid, albeit different pictures of the city. Thus, while examining textual tone furthers our understanding of disclosure, it is limited to a particular perspective, which does not take into account all the channels through which meaning is conveyed.

The application of content analysis, in general, and textual tone analysis, in particular, has gained increased attention in disclosure research in recent years. Davis et al. (2008), Demers and Vega (2008), Henry (2008), and Sadique (2008) extract the tone of the wording of quarterly earnings press releases and relate it to things such as stock returns, volatility, and firm performance. All of these textual analysis studies find statistical significance for the linguistic tone of disclosure documents, suggesting that relevant information is conveyed by managers in their word choices. However, none of these studies examine the additional disclosures embedded in the language of the earnings conference calls. Li (2008) acknowledges that managers face more constraints when communicating with investors through formal reports and announcements (e.g. annual reports, earnings announcements, etc.) and, consequently, suggests that conference calls may provide a better setting in which to explore the relation between linguistic content and firm performance. In other words, quarterly earnings conference calls provide a forum in which to more fully examine corporate disclosure as executives interact with call participants through the unscripted question and answer sessions.

We test the information asymmetry between managers and investors using conference call textual tone. In order to examine the relation between conference call disclosure, as captured by call tone, and stock returns, we employ computer based content analysis to quantify the linguistic tone of quarterly earnings conference calls. More specifically, we look at the effect of call content on investors by examining the relation between the linguistic tone of call transcripts and both earnings announcement abnormal returns and the post-earnings-announcement drift.

The post-earnings-announcement drift,[2] the phenomenon of abnormally high (low) returns following positive (negative) earnings announcements, has puzzled financial market researchers for over four decades (Ball and Brown, 1968; Bernard and Thomas, 1989, 1990). However, recent studies contend that the drift may be due to uncertainty (Lewellen and Shanken, 2002; Brav and Heaton, 2002; Liang, 2003; Zhang, 2006; and Francis et al., 2007). They argue that in cases of greater uncertainty, where investors must learn about cash flows and future firm prospects, information may be more slowly incorporated into asset prices resulting in the empirical manifestation of the drift. While the textual tone may provide more precise information to investors, the drift may still be directly related to the tone. A delayed price reaction may occur if some investors are slow to receive or process the tone information. For example, many investors may be unable to listen to conference calls and, instead, view them on firm websites after the call.

While the above studies use various proxies to measure uncertainty, no universally accepted uncertainty measure has been established in the literature (Zhang, 2006). We perform our analysis using an uncertainty indicator that is closely related to the certainty of investor cash flows by separately considering dividend paying firms and those that do not pay dividends. Several recent studies document that firms who pay dividends tend to be larger, more profitable firms whose earnings are much more certain than small, growth oriented firms that do not pay dividends (DeAngelo et al., 2004; Denis and Osobov, 2008). Additionally, Hoberg and Prabhala (2009) show that the propensity to pay dividends is inversely related to firm risk. Thus, in light of the uncertainty hypothesis for the post-earnings-announcement drift (Lewellen and Shanken, 2002; Brav and Heaton, 2002; Liang, 2003; Zhang, 2006; Francis et al., 2007), the uncertainty differences between dividend and non-dividend paying firms may lead to a better understanding of the importance investors place on various types of earnings signals. Specifically, we expect conference call tone to matter more for firms with greater investor cash flow uncertainty (non-dividend payers).

For the most part, research has maintained an exclusive focus on the numerical value of earnings when trying to explain the post-earnings-announcement drift anomaly. We focus our analysis on the information found in the text of quarterly conference calls to explore the relation between the linguistic tone, or sentiment, of the conference call words and both the initial market reaction and the post-earnings-announcement drift. We do so using a hand collected sample of over 2800 conference call transcripts spanning 16 consecutive quarters in the post Regulation Fair Disclosure (FD) era. We differentiate between the two major components of an earnings conference call by separating management's prepared remarks and the more spontaneous question and answer portion of the call. This allows us to control for the information conveyed in the written press release to determine the incremental disclosure found in the conference call dialogue between management and analysts. Accordingly, we expect a significant relation between the tone of the question and answer portion of the conference calls and stock returns. Furthermore, if the post-earnings-announcement drift is rooted in uncertainty, we expect this relation to be more pronounced during the drift period for firms that do not pay dividends when compared to dividend paying firms.[3]

Following Tetlock (2007), we employ textual analytical techniques using the General Inquirer package.[4] However, unlike Tetlock, who relies on the program's default dictionary (Harvard IV-4 Psychosocial) to isolate linguistic tone, we also incorporate a custom dictionary with earnings specific language. Using the two separate dictionaries allows us to conduct an out-of-sample examination of the argument by Loughran and McDonald (2011) that custom, context specific dictionaries are most appropriate for use in financial research. While Loughran and McDonald analyze printed 10-K disclosures, they indicate that whether their results hold for samples of other types of financial text is an important question.

[2] Also referred to as simply ''the drift''.

[3] This study is largely limited to examining the incremental disclosures reflected in the tone of the question and answer portion of the conference calls. However, we are able to control for the earnings surprise, the tone of the prepared statements by management (which proxy for the tone of the press release), the length of both the prepared statements and question and answer portions of the call using word counts, and numerous additional ''standard'' controls in order to isolate the effects of call tone.

[4] While many text based content analysis programs exist, thus far the General Inquirer has been the most widely used application in the burgeoning linguistic analysis finance literature since many studies follow Tetlock (2007) and Tetlock et al. (2008).

We add to the textual analysis literature by finding out-of-sample support for the assertion of Loughran and McDonald (2011) in that the custom, earnings specific dictionary is much more powerful in detecting relevant conference call tone. We also extend the literatures that deal with conference call disclosures and abnormal returns. We find that conference call discussion tone has highly significant explanatory power for initial reaction window abnormal returns as well as the post-earnings-announcement drift. This holds after controlling for both the magnitude of the earnings surprise and the tone of the prepared managerial statements. Furthermore, conference call tone is better able to predict cumulative abnormal returns during the drift window than the numerical value of unexpected earnings.

We find dramatic differences in the predictive ability of call tone between dividend payers and non-payers in firm-level and portfolio-level analyses. Providing support for the uncertainty explanation for the drift, conference call question and answer tone matters more when firms do not pay dividends. In other words, nonfinancial information is relied on more heavily by investors in cases of greater investor cash flow uncertainty, consistent with Amir and Lev (1996).

Overall, this study makes three distinct contributions to the broader disclosure literature by way of studying the effects of the textual tone of disclosure and textual tone within a less formal disclosure arrangement. First, we examine the question and answer session within a conference call. Second, we look at the effects of disclosure through the measurement of the post-earnings-announcement drift. Third, we demonstrate a simple, effective measure of textual tone.

The remainder of this study is developed in the following sections. We review the literature in the next section. The third section describes our sample. In the fourth section we outline the content analytical techniques and empirical methods. The fifth section provides results. The final section concludes.

2. Literature review

Beaver (1968) shows that capital markets react to the information content in earnings reports. Ball and Brown (1968) find evidence that stock returns continue to drift in the same direction as unexpected earnings. That is, stock returns tend to drift upward (downward) after earnings announcements with positive (negative) earnings surprises. This post-earnings-announcement drift phenomenon stands in stark contrast to the concept of efficient markets, where the information content of the numerical value of earnings should be fully incorporated into stock prices within a short period of time after the information becomes publicly known.

Researchers have been unable to fully attribute the drift to risk adjustments, market frictions, or the psychological biases of market participants (Bernard and Thomas, 1989, 1990; Bernard, 1992). Bernard and Thomas argue that the evidence is consistent with a delayed response explanation for the drift, where the market fails to immediately incorporate the information in the earnings surprise signal. Abarbanell and Bernard (1992) indicate that the drift may be rooted in the failure of market participants to fully understand the information content of current earnings. Such information is not limited to an exclusive focus on the numerical value of the earnings alone, but includes statements that typically accompany the earnings announcement. These statements provide context and add to our understanding of firm performance.[5]

To effectively analyze such communication, a technique widely used across numerous disciplines, called content analysis, helps us better understand the relation between the wording in statements and stock prices.[6] Content analysis is defined as the analysis of the manifest and latent content of a body of communicated material through classification, tabulation, and evaluation of its key symbols and themes in order to ascertain its meaning and probable effect.[7] As applied in this study, this involves the quantification of text based information through word classification and tabulation.

Recent finance studies which rigorously incorporate text based content analysis include Tetlock (2007) and Tetlock et al. (2008). Both studies capture the sentiment of the media by applying content analysis to articles found in widely disseminated media sources. Tetlock (2007) analyzes the daily Wall Street Journal column, *Abreast of The Market*, and finds that high media pessimism predicts downward pressure on market prices and subsequent reversion. Tetlock et al. (2008) examine the relation between simple quantitative language measures and firm fundamentals. They find that the negative words in Wall Street Journal and Dow Jones News Service stories about individual S&P 500 firms can be used to predict individual firms' accounting earnings and stock returns.

Engelberg (2008) counts the number of negative words as defined by the Harvard IV-4 Psychosocial Dictionary in the text of Dow Jones News Service stories about firms' earnings announcements following the method outlined in Tetlock (2007). In his attempt to link the qualitative earnings information embedded in the financial media with the post-earnings-announcement drift, Engelberg finds that text based "soft" information has additional predictability for equity prices beyond the "hard" information of the numerical earnings surprise. Specifically, the linguistic content of the news stories predicts larger price changes at longer horizons than unexpected earnings, suggesting that frictions in information processing may generate the drift. However, Engelberg does not consider more direct sources of information such as earnings announcements or the corresponding earnings conference calls.

Several recent studies incorporate content analysis to examine the information content of the wording of earnings announcements (Davis et al., 2008; Demers and Vega, 2008; Henry, 2008; Sadique, 2008). Davis et al. (2008) examine whether managers use linguistic style in earnings press releases to provide information about expected future firm performance using a broad sample of firms. They find a positive association between optimistic tone in the press release and future return-on-assets. Using a sample in the telecommunications and computer industry, Henry (2008) measures the relative proportion of positive and negative words in earnings press releases to capture the statements' tone. She shows that tone affects the market's initial reaction to the earnings announcements.

Sadique (2008) extend this work by analyzing the content of both earnings announcements and related press coverage in relation to stock returns and volatility. They find that positive tone is related to increased returns and decreased volatility in a sample of S&P 100 firms. Demers and Vega (2008) examine whether the "soft" information in earnings press releases is incrementally informative over a company's reported "hard" earnings news. They find that managerial word choice is related to both abnormal returns and idiosyncratic volatility for up to 60 days following the announcement. Demers and Vega argue that it takes longer for the market to understand the implications of the wording of the earnings announcements relative to the numerical value of the

[5] We do not attempt to directly test between rational and behavioral explanations for the post-earnings-announcement drift in this study. However, since the social communications literature finds that people are affected by the words of others (see Petty and Cacioppo, 1986), further research which attempts to distinguish between rational and behavioral causes of the drift may benefit from the incorporation of textual tone analysis.

[6] Text based content analysis techniques, developed decades ago, were originally applied in journalism, psychology, communications, and other social sciences. See Krippendorf (2004) for a complete history of content analysis, including its development and application across disciplines.

[7] http://www.merriam-webster.com/dictionary/content%20analysis.

earnings surprise and that the market response varies by firm characteristics.

Overall, these studies (Davis et al., 2008; Demers and Vega, 2008; Henry, 2008; Sadique, 2008) suggest that the information content of the words accompanying an earnings announcement provides additional insight beyond that which we learn by scrutinizing the numerical value of the earnings surprise in isolation. However, none of these studies analyze earnings conference calls, an information rich source with potentially greater information content than in the more narrow earnings press release.

Frankel et al. (1999) note that conference calls are often used to supplement quarterly earnings releases and argue that managers use the calls to explain the implications of unusual or extraordinary items. This, in turn, aids analysts in determining the extent to which earnings surprises are permanent or temporary. Frankel et al. (1999) find that conference calls provide information to market participants over and above the information contained in the corresponding press release. Elevated return volatilities and trading volume during the conference calls suggests that large investors trade in real time based on the information released during the calls.

Sunder (2002) shows that firms disclose value relevant information to the public during conference calls following the passage of Regulation FD.[8] Irani (2004) indicates that conference calls are a medium increasingly used by firms to disclose information to market participants and shows that post-FD conference calls are more informative than pre-FD conference calls. The National Investor Relations Institute indicates that next to news releases to wire services, conference calls are the most widely used means for disseminating corporate information to the investment community (NIRI, 2004).[9] They further state that conference calls are often used as a forum in which firms provide detailed information.

Kimbrough (2005) explains that conference calls typically include opening remarks by management, which generally reiterate the press release, followed by a question and answer session where details not contained in the press releases are often disclosed. He notes that conference call discussions provide managers with an opportunity to comment on recent results and emphasize their implications for future financial performance. Kimbrough finds that the initiation of conference calls is associated with a significant reduction in the post-earnings-announcement drift, suggesting that conference calls improve the efficiency of the market reaction to earnings announcements.

A few recent studies incorporate the linguistic information content of earnings conference calls in their analyses. Frankel et al. (2010) use conference call tone as a proxy for the prevailing state of relations between a firm and its investors in their examination of investor relations costs. They find a positive relation between conference calls returns and linguistic tone. Mayew and Venkatachalam (forthcoming) analyze conference call audio files using vocal emotion analysis software. They examine the relation between managerial affective states (voice inflection), firm fundamentals, and performance. The number of negative words in each call is used as a control variable. Neither Frankel et al. nor Mayew and Venkatachalam differentiate between the wording of the introduction and the question and answer portion of the calls.

Unlike these conference call studies, we focus our analysis on the information found in the text of quarterly conference calls to explore the relation between the linguistic tone of the conference call dialogue and both the contemporaneous market reaction and the post-earnings-announcement drift. We differentiate between the two major components of an earnings conference call by separating management's prepared remarks and the unscripted question and answer portion of the call. This distinction is of theoretical importance.

Dye (2001) claims that the theory of voluntary disclosures is a special case of game theory where firms will only disclose information that is favorable to the firm, and will not make disclosures that are unfavorable to the firm. Consequently, the interpretation of the information, both stated and unstated, should be made with the understanding that the firm will behave accordingly. Healy and Palepu (2001) describe potential incentives for managers to make self-serving voluntary disclosures, rendering the credibility of the sentiment in the prepared statements to be suspect. However, the Q&A portion of the call includes questions and statements from third-party intermediaries, which provides more assurance about the quality of the overall Q&A tone signal.

Healy and Palepu also document the general assumption that managers have superior information to outside investors regarding the expected future performance of their firms. Verrecchia (2001) suggests that information asymmetry reduction is an area with great potential for establishing a link between disclosure and its economic consequences. We test this potential information asymmetry between managers and shareholders using tone oriented content analysis to investigate the incremental information content of the Q&A component of quarterly earnings conference calls, above and beyond that found in the prepared statements of management and the earnings surprise itself.[10] Moreover, we determine if the market responds to this information efficiently by relating it to short- and medium-term stock returns.

Furthermore, unlike other conference call studies, we examine differences across firms based on their level of cash flow uncertainty. The role of uncertainty in asset pricing has received increasing attention in the literature in recent years.[11] We use dividends as an indicator of relative certainty where the dividend paying distinction enables us to check for differences between firms based on the certainty of the cash flows to investors.[12] DeAngelo et al. (2004) and Denis and Osobov (2008) document that dividend payout has become concentrated among firms who generate substantial and stable earnings. Hoberg and Prabhala (2009) find that risk is negatively related to the probability of being a dividend payer. They show that a one standard deviation change in idiosyncratic risk shifts the probability of being a dividend payer by 39%.

Henry (2008) deliberately chooses a sample of firms drawn from the telecommunications and computer services industries for the period 1998–2002 because of the inherent uncertainty in both the industry and time period. She argues that, according to Amir and Lev (1996), non-financial information is relied on by investors more heavily during periods of uncertainty. Dividend payout differentiation allows us to examine this conjecture across a broad sample by comparing the non-numerical information content of earnings conference calls between firms with greater and lesser cash flow uncertainty.

8 The SEC rule commonly known as "Regulation FD" opened up earnings conference calls to the public, effective October 23, 2000. The final SEC rule is available at http://www.sec.gov/rules/final/33-7881.htm.

9 This is also noted in Frankel et al. (1999) using a NIRI (1996) report.

10 We recognize that releasing information publicly through conference call disclosures may preempt some future private information events, which potentially reduces information asymmetry between investors (Brown et al., 2004). However, in the context of our analysis, we view this as a down-stream benefit whose origin lies in the more direct reduction of asymmetry between management and investors.

11 For example, Lewellen and Shanken (2002), Brav and Heaton (2002), Liang (2003), Zhang (2006), and Francis et al. (2007) all focus on uncertainty in asset pricing from either a theoretical or empirical perspective.

12 Although the relevance of dividend payout has been the subject of much study, and controversy, over the past fifty years (Miller and Modigliani, 1961), recent theoretical, empirical, and survey evidence suggests that dividend payout is a nontrivial matter that has implications for firm value (Allen and Michaely, 2002; Brav et al., 2005; DeAngelo and DeAngelo, 2006).

Table 1
Uncertainty proxy correlations.

	DIV	COUNT	SIZE	BM	ROA	LEVERAGE	VOLUME	VOLATILITY	ANALYST	SPREAD
DIV	1.00									
COUNT	−0.07	1.00								
SIZE	0.10	0.44	1.00							
BM	−0.02	0.00	−0.12	1.00						
ROA	0.14	0.08	0.23	−0.05	1.00					
LEVERAGE	0.18	0.05	0.15	0.02	−0.09	1.00				
VOLUME	−0.05	0.47	0.77	−0.09	0.11	0.14	1.00			
VOLATILITY	−0.25	−0.16	−0.48	0.01	−0.21	−0.16	−0.20	1.00		
ANALYST	−0.06	0.33	0.52	−0.07	0.13	−0.04	0.48	−0.23	1.00	
SPREAD	−0.10	−0.20	−0.41	0.06	−0.15	−0.04	−0.31	0.33	−0.26	1.00

This table provides correlations for various uncertainty proxies for the full sample of 2880 conference calls. DIV is an indicator variable equal to 1 if the firm pays dividends, and zero otherwise. COUNT is the number of words, in thousands, for a given conference call. SIZE is the log of firm market capitalization from the previous quarter. BM is the ratio of book-to-market equity as of the end of the previous quarter. ROA is return on assets, in percent, calculated as net income divided by total assets multiplied by one hundred. LEVERAGE is expressed in percent as the ratio of total liabilities to total assets multiplied by one hundred. VOLUME is the log of the total share trading volume on day zero. VOLATILITY is in percent and is calculated as the standard deviation of daily returns for the ninety trading-day period ending 10 days prior to the conference call multiplied by one hundred. ANALYST is the log of the number of analysts who cover a given firm. SPREAD is the bid-ask spread on the day of the conference call calculated as ask minus bid, scaled by stock price.

3. Data

This study covers 16 consecutive quarters during the post Regulation FD 4-year period 2004–2007. Prior to this time, conference call transcript availability is limited.[13] After this time, markets experienced a precipitous decline initiated by the bursting of the housing bubble. Calls held during the sample period discuss firm performance starting with the fourth quarter of 2003 and run through the third quarter of 2007 due to the several week lag between the close of each quarter and the corresponding earnings announcements and conference calls.

We use a pseudo-random sample comprised of 2880 observations. REITs, ADRs, closed end funds, and units are excluded by limiting CRSP data to share codes 10, 11, and 12. Other financials (SIC 6000-6999) and utilities (SIC 4900-4949) are excluded as well. Prior to random selection, we drop all firms which fail to meet data sufficiency requirements. Firms need at least 60 days of daily returns on CRSP prior to the conference call and 60 days of daily returns following the call in order to estimate abnormal returns and show the drift (Bernard and Thomas, 1989). In addition, seasonally lagged earnings-per-share (EPS) data obtained from Compustat are necessary to calculate the earnings surprise.

For each quarter we sequentially sort the universe of merged CRSP and Compustat firms into terciles by size[14] and book-to-market equity (BM)[15] in order to ensure that our sample contains a cross section of firms with varying characteristics.[16] Size is the market capitalization of each firm calculated as the number of shares outstanding times the market price of the stock at the end of the preceding quarter. BM is calculated as the two-fiscal-quarter lagged Compustat book value of equity divided by the market capitalization of each firm at the end of the preceding quarter.

Table 2
Earnings announcement and conference call timing.

Number of trading days after announcement	Number of calls	Percent of observations	Cumulative percent
0	2423	84.13	84.13
1	439	15.24	99.38
2	6	0.21	99.58
3	7	0.24	99.83
4	3	0.10	99.93
5	1	0.03	99.97
6	1	0.03	100
Total	2880	100	

This table shows the timing relation between earnings announcements and conference calls. Most of the calls take place on either the same day as the announcement or on the following day.

Within the nine size-BM portfolios we further separate dividend and non-dividend paying firms, resulting in 18 characteristic portfolios for each quarter. We use dividends as a proxy for cash flow uncertainty. Table 1 reports the correlations between dividends and measures of firm uncertainty. Although this is an imperfect measure, firms that pay dividends tend to be larger firms, are less volatile, have higher ROAs, and have lower bid-ask spreads, which are characteristics consistent with more certain future cash flows. Additionally, the length of the conference call is shorter, which along with the other proxies for uncertainty around the earnings announcement, suggests that separating firms by dividends is a good proxy for investor insecurity about future returns. Ten firms are then randomly selected from each of the characteristic portfolios for each of the 16 quarters, resulting in ($18 \times 16 \times 10$ =) 2880 firm-quarter observations.

Conference call transcripts are obtained from two sources, Fair Disclosure Wire and The American Intelligence Wire. Call dates generally accompany the transcripts and are confirmed using Thomson Streetevents. Table 2 shows the relative timing of earnings announcements and conference calls for our sample of 2880 firm-quarter observations.[17] For the number of trading days after announcement, a zero indicates that a conference call is held on the same day as the earnings announcement. Likewise, a one indicates that a conference call is held 1 day after the earnings announcement date, a two indicates that a call takes place two trading days after the announcement date, and so forth. At 84%, the vast majority of calls take place on the same day as the earnings announcement. Over 15% of

[13] While Regulation FD took effect at the end of 2000, it took a few years before electronic conference call transcripts became widely available. Starting the sample earlier would have resulted in a more pronounced bias towards large, well established firms since they are more likely to receive greater attention and merit transcription coverage. We acknowledge that even during the 2004–2007 period there may be some bias inherent in each of our firm-quarter observations since we require each observation to have an electronically available transcript. We attempt to mitigate this to some extent by drawing a pseudo-random sample wherein we force the sample to have an equal number of small, medium, and large firms.

[14] Size sort based on NYSE market capitalization breakpoints obtained from the website of Ken French. Since the breakpoints are only available at five percent increments, terciles are extrapolated as follows: $((p35 - p30)/5 * 3 = p33)$ and $((p70 - p65)/5 * 2 = p67)$.

[15] Negative BM equity firms are dropped.

[16] As such, we avoid the potential biases of having a sample skewed towards either large or small firms, or value or growth firms.

[17] Earnings announcement dates are obtained from the Compustat Quarterly database, item "RDQ – Report Date of Quarterly Earnings".

the calls are held on the following day.[18] The remaining calls (0.6%) are all held within the next few days.

4. Content analysis techniques and empirical methods

Nearly 50 years ago Harvard researcher Philip J. Stone created the first content analysis computer program, the General Inquirer (GI).[19] Since that time, computer based content analysis has been widely used across many disciplines to quantify text based information in an effort to better understand the text's meaning. To accomplish this in a rigorous way, content analysis software applies predefined rules which aid in word recognition, categorization, and tabulation.

With continued refinement over the years, the GI has maintained its place as a powerful tool for quantifying linguistic information. Under its current iteration, the GI processes text by first applying over 6300 disambiguation rules as part of the word recognition process. Following word recognition, the program applies a dictionary that categorizes each word according to its meaning and then tabulates a raw count with the number of words that fall within each predefined category. By default, the GI applies the Harvard IV-4 Psychosocial Dictionary which contains approximately 12,000 words and 77 separate categories within which a word can be classified.[20] It is also possible to create custom dictionaries with categories and word lists that are specifically tailored to the context of interest. Other computer based textual analysis programs operate in a similar fashion, however none of them currently have the capacity to handle a dictionary as large and robust as the Harvard dictionary.

In recent work, Tetlock (2007), Tetlock et al. (2008), Engelberg (2008), and Frankel et al. (2010) all utilize the GI and the Harvard dictionary to quantify the linguistic tone of written material. Davis et al. (2008), Henry (2008), and Sadique (2008) use the Diction software package and a custom dictionary as found in Henry (2008). Demers and Vega (2008) use both the GI and Diction in their linguistic analysis and rely on the "off the shelf" dictionaries that each of the programs uses by default. In comparing the two packages, Demers and Vega find the default GI word list (Harvard dictionary) to be more powerful in predicting stock returns.

While textual analysis techniques are well established and have been widely used across many disciplines such as journalism, communications, psychology, sociology and other social sciences, their application to financial research is still in its infancy. The question of interdisciplinary dictionary transferability comes into play. In other words, is it appropriate to apply a dictionary that is commonly used in psychological research in a financial context? Henry (2008) and Loughran and McDonald (2011) both contend that each discipline has its own dialect, where words take on specific meanings in specific contexts which may not translate effectively across disciplines. Loughran and McDonald examine the Harvard dictionary in relation to a dictionary of their own creation using corporate 10-K filings and find that the Harvard dictionary often misclassifies terms in annual reports. That is, words which take on a negative meaning in every day speech do not necessarily have a negative meaning in the context of corporate annual reports.

Moreover, textual analysis techniques have been criticized as being atheoretical (Berelson, 1952; West and Fuller, 2001). General purpose dictionaries, in particular, have been subject to this criticism, while special-purpose (context-specific) dictionaries are more easily supported within a theoretical superstructure (Psathas, 1969; West and Fuller, 2001). Using textual tone in communications research has its theoretical basis in the idea that extracting tone is the same process that people undergo in the construction of meaning (Laffal, 1970; West and Fuller, 2001). Hart (2001) argues that psychotherapists, literary critics, and others, instinctively sense subtle differences in tone when words are used in different settings or contexts. Thus, the specific dictionary used to extract linguistic tone from a body of text is of theoretical importance in providing a vital link between the formulation of the problem and the mechanics of analysis (Holsti, 1963; Diefenbach, 2001).

In order to provide an out-of-sample examination of dictionary power, we use both the Harvard dictionary and a custom, finance oriented, dictionary. Unlike previous studies that use the Harvard dictionary and only incorporate two word categories, positive and negative, we use additional categories that contain words that would generally be considered either positive or negative in nature. Specifically, we employ eight of the Harvard dictionary categories that are relevant to this study including positive, negative, active, passive, strong, weak, overstatement, and understatement and condense them into a single measure of conference call tone. For the custom dictionary we use the word lists defined in Henry (2008). The Henry dictionary is a simple list of "domain relevant" words that are categorized as either positive or negative. The custom dictionary was created specifically using quarterly earnings press releases which is appropriate for our examination of quarterly earnings conference calls.[21]

We use the GI for word recognition and apply both the Harvard and Henry dictionaries for comparison. After processing the sample of conference call text files through the content analysis software, raw word counts are divided by the total words in their respective conference calls to produce a relative measure ($REL\_M_{i,j}$) for each category i and conference call j:

$$REL\_M_{i,j} = \frac{Raw\ Word\ Countfor\ Category_{i,j}}{Total\ Call\ Word\ Count_j} \tag{1}$$

where the $REL\_M_{i,j}$ categories are $Positive_j$, $Negative_j$, $Active_j$, $Passive_j$, $Strong_j$, $Weak_j$, $Overstated_j$, and $Understated_j$.

The eight relevant Harvard dictionary categories comprise four opposing pairs. Expressed in ratio form we can tell, for example, how positive a call is relative to how negative it is. To arrive at an overall numerical representation of conference call tone for each conference call j ($TONE_{1,j}$), we take the first principal component of the four relevant category ratios:

$$TONE_{1,j} = First\ Principal\ Component \left\{\frac{Positive_j}{Negative_j}, \frac{Active_j}{Passive_j}, \frac{Strong_j}{Weak_j}, \frac{Overstated_j}{Understated_j}\right\} \tag{2}$$

This measure is used in place of simply isolating the positive-to-negative ratio since investors will likely see words that fall within the active (passive), strong (weak), or overstated (understated) categories as "positive" ("negative"). Taking the first principal component allows collapsing of the variation across the four ratios into one simple measure of overall conference call tone. Thus, we treat

[18] For the 439 calls that fall within the "one trading day after announcement" category, the earnings announcements were typically released after hours on the previous day.

[19] Stone et al. (1966) describe the program and its potential application across numerous disciplines. For many years, the software application needed to be installed on individual computers. More recently, GI has transitioned to a web-based platform.

[20] For completeness, we note that the GI outputs a total of 182 separate categories since it also incorporates the Lasswell general purpose dictionary by default.

[21] See the Appendix for further detail including a description of the categories and word lists. We use the Henry dictionary instead of the negative word list created by Loughran and McDonald (2011) because it is both more concise and specifically relevant to the context of our examination of earnings conference calls. Loughran and McDonald's list was created using 10-K filings, while Henry's list was created using earnings announcements. However, we feel the Loughran/McDonald word list would likely produce results similar to the Henry dictionary since they are both financially oriented.

the *TONE* variable as a metric of relative positivity.

For robustness, we also construct the tone variable using an alternative form of the Harvard dictionary category ratios described above. First we take the difference in the opposing $REL\_M_{i,j}$ categories and divide by the sum of the two. Specifically, for each conference call j we calculate the following four ratios using the $REL\_M_{i,j}$ categories:

$$PN_j = \frac{Positive_j - Negative_j}{Positive_j + Negative_j} \tag{3}$$

$$AP_j = \frac{Active_j - Passive_j}{Active_j + Passive_j} \tag{4}$$

$$SW_j = \frac{Strong_j - Weak_j}{Strong_j + Weak_j} \tag{5}$$

$$OU_j = \frac{Overstated_j - Understated_j}{Overstated_j + Understated_j} \tag{6}$$

In this form the ratios are bounded between −1 and 1, consistent with the tone measures of Uang et al. (2006) and Henry (2008). For the second overall numerical representation of conference call tone for each conference call j ($TONE_{2,j}$), we take the first principal component of the four relevant category ratios:

$$TONE_{2,j} = First\ Principal\ Component\{PN_j, AP_j, SW_j, OU_j\} \tag{7}$$

For the Henry measure for each conference call j ($H\text{-}TONE_j$), we simply express the $REL\_M_{i,j}$ categories $Positive_j$ and $Negative_j$, as defined in the earnings specific Henry dictionary, in the two ratio forms described previously:

$$H\text{-}TONE_{1,j} = \frac{Positive_j}{Negative_j} \tag{8}$$

and

$$H\text{-}TONE_{2,j} = \frac{Positive_j - Negative_j}{Positive_j + Negative_j} \tag{9}$$

We calculate cumulative abnormal returns to isolate the stock price reaction to the earnings announcements and corresponding conference calls. Daily abnormal returns are defined as size adjusted returns calculated as the difference between the raw return for stock j on day t and the mean return of a portfolio of all firms in the same size decile (Foster et al., 1984; Bernard and Thomas, 1989, 1990)[22]:

$$AR_{j,t} = R_{j,t} - R_{p,t} \tag{10}$$

where $AR_{j,t}$ is the abnormal return for firm j on day t, $R_{j,t}$ is the return for firm j on day t, and $R_{p,t}$ is the mean return on day t for all firms in the same size decile as firm j.

We cumulate abnormal returns over two different windows, days minus one through one and days 2 through 60, where day zero is the date of the quarterly earnings conference call. We analyze the effect of conference call tone on the contemporaneous stock price reaction using CARs from the initial 3-day period, while the CARs from the longer period allow us to examine the relation between tone and the post-earnings-announcement drift.[23] While Bernard and Thomas (1990) show evidence of the drift through three quarters for a sample of earnings announcements from 1974 to 1986, others find that the magnitude, and perhaps duration, of the drift has declined over time. Campbell et al. (2009) find that the drift survives over the 60 days following the announcement in their 1995–2000 sample period, but that the effect is weaker than in Bernard and Thomas. Thus, we use two cumulative abnormal returns measures for each stock associated with conference call j, $CAR(-1,1)_j$ and $CAR(2,60)_j$, where:

$$CAR(-1,1)_j = \sum_{t=-1}^{1} AR_{j,t} \tag{11}$$

$$CAR(2,60)_j = \sum_{t=2}^{60} AR_{j,t} \tag{12}$$

After calculating the cumulative abnormal returns measures we conduct a preliminary analysis by sorting the sample into tone quintiles and check for differences in CARs between the high and low tone portfolios. We compare the CAR means and medians of the high and low tone quintiles using t-tests and Wilcoxon rank-sum tests, respectively. The comparison is done for all four tone measures and both series of estimated CARs, which enables us to check the power of the tone measures and the two different dictionaries used to construct them. To further illustrate the differences across tone portfolios, we plot the cumulative abnormal returns from 10 days prior to the conference call through 60 days after the conference call for firms in the high tone portfolio, firms in the low tone portfolio, and an average of all firms in the sample across the tone quintiles.

Unexpected earnings, $SURP_j$, are calculated using a seasonal random walk model where the difference between the earnings-per-share and the earnings-per-share in the same quarter of the prior year is scaled by the stock price at the close of the lagged quarter:

$$SURP_j = \frac{(EPS_j - EPS_{j,q-4})}{StockPrice_{j,q-4}} \tag{13}$$

The seasonal measure of earnings surprise is widely used in the literature along with a measure calculated as the stock price scaled difference in actual and mean analyst forecast earnings. However, the literature is inconclusive with respect to which measure is most appropriate.[24] Recent research shows that the analyst forecast based measure suffers from the finding that analysts only forecast GAAP earnings 65% of the time. In other words, 35% of the time analysts forecast an earnings number that does not correspond to the actual GAAP earnings number used in calculating GAAP earnings surprises (Doyle et al., 2003; Collins et al., 2009). Sadique (2008) argue that analysts forecast *pro forma* earnings.

Furthermore, survey evidence shows that managers place the greatest weight on four quarter lagged earnings per share as their earnings benchmark. Graham et al. (2005) find that 85.1% of CFO's consider earnings in the same quarter of the prior year to be the most important earnings benchmark, followed secondly by mean analyst forecast estimates. Thus, using the seasonal measure may provide the most appropriate context in which to examine the relation between earnings and conference call tone since managerial tone may be more directly related to their performance relative to the benchmark they consider to be most important. Regardless, we re-

[22] Size deciles are determined using NYSE size (market capitalization) breakpoints as found on Ken French's website. While the literature shows that short window cumulative abnormal returns are unaffected by risk adjustment, we find that our sample is also mostly unaffected by size adjustment. We obtain nearly identical results using simple Market Adjusted Returns (MAR) without size adjustment and the CRSP value-weighted index as the market proxy.

[23] Davis et al. (2008), Demers and Vega (2008), Engelberg (2008), and Henry (2008) all use three-day CARs to examine the initial reaction period. Foster et al. (1984), Bernard and Thomas (1989), and many subsequent studies use CARs of roughly 60 trading days to examine the post-earnings-announcement drift.

[24] For example, Akbas et al. (2010) only use the seasonal measure. Sadique (2008) use both measures in their analysis and find no substantive differences between them. However, they focus on, and only tabulate results for, the seasonal measure. Henry (2008) also uses the seasonal measure to control for unexpected earnings, although she incorporates an indicator variable set to one if earnings are greater than the consensus analyst forecast as well. Davis et al. (2008) and Engelberg (2008) both use an analyst forecast method exclusively. However, Engelberg notes that his sample is biased against small firms since analysts are more likely to cover large firms. Brown et al. (1987) show that both the seasonal and analyst forecast based earnings surprise proxies are prone to measurement error.

port our results using the seasonal measure to prevent substantial attrition in our hand collected sample.[25]

To demonstrate performance variation in the cross section of earnings surprises, we sort the sample into SURP quintiles and visually check for differences in CARs between the high and low SURP portfolios. In the same manner as the tone based illustration described above, we plot the cumulative abnormal returns from 10 days prior to the conference call through 60 days after the conference call for firms in the high SURP portfolio, firms in the low SURP portfolio, and an average of all firms in the sample across the SURP quintiles.

We also visually check the relation between abnormal returns and conference call tone while controlling for SURP by conducting two-way sequential sorts on tone and earnings surprise. For each dimension of the sort we form terciles that contain the mean CAR for their respective portfolios and plot the results. When holding the earnings surprise constant, we expect to see CAR differences in high and low tone portfolios.

At the firm level, we examine the effect of conference call tone on stock returns, while controlling for unexpected earnings, by running cross-sectional regressions of the following form:

$$CAR_j = \gamma_{0,i} + \gamma_{1,i}SURP_{i,j} + \gamma_{2,i}TONE_{i,j} + CONTROLS_j + \varepsilon_j \tag{14}$$

where CAR_j represent the cumulative abnormal return measures for conference call j defined above; $TONE_{i,j}$ is interchangeable with $H-TONE_{i,j}$, both as defined above and where $i = 1$ or 2; and $CONTROLS_j$ includes measures of call length, firm size, book-to-market equity, profitability, leverage, trading volume, returns volatility, and analyst coverage for firm j.[26] Call length provides an indication of the quantity of information contained in the conference call and is measured for the two separate call components as their respective word counts, in thousands (INTRO COUNT and Q&A COUNT). SIZE is the log of firm market capitalization at the end of the previous quarter. The ratio of book-to-market equity (BM) as of the end of the previous quarter picks up disclosure differences associated with growth opportunities. The percent return on assets (ROA), calculated as net income divided by total assets multiplied by one hundred, accounts for increased investor information demand when firms are not profitable. Similarly, the ratio of total liabilities to total assets, scaled by one hundred, (LEVERAGE) controls for increased information demand when firms are experiencing financial distress. VOLUME is the log of the total share trading volume on the day of the conference call where trading volume for NASDAQ firms is divided by two to avoid double counting. VOLATILITY, in percent, is calculated as the standard deviation of daily returns for the 90 trading day period ending 10 days prior to the conference call scaled by one hundred, since investors may require more information during periods of greater uncertainty. The data for SIZE, BM, ROA, LEVERAGE, VOLUME, and VOLATILITY are from CRSP and Compustat. Analyst coverage data (ANALYST) are from Thomson Reuters' First Call Historical Database and calculated as the log of the number of analysts covering a given firm immediately prior to the announcement. DECLARATION is an indicator variable set equal to one if a firm declares dividends within the 3-day window around the conference call $(-1, 1)$, and zero otherwise. Standard errors are corrected for potential heteroscedasticity following White (1980). We also adjust the standard errors for time- and firm-effects by clustering by firm and by quarter using the methodology outlined in Petersen (2009).[27]

We expect the earnings surprise control variable to be the dominant predictor of abnormal returns during the initial reaction period ($t = -1$ to $t = 1$) and anticipate the coefficient will be positive and significant. Similarly, if additional information is disclosed during the conference calls, as proxied by call tone, we expect the tone measures to be both significant and positively related to the initial reaction window CARs. The efficient markets hypothesis suggests that the full impact of both the earnings surprise and the tone measures should be completely incorporated into the initial response. However, the coefficients on both the earnings surprise measure and the tone measures may be significant over a longer horizon ($t = 2$ to $t = 60$) given the pervasive post-earnings-announcement drift. Engelberg (2008) and Demers and Vega (2008) suggest that soft information, such as our tone measures, has greater predictability for returns at longer horizons. In other words, the market may impound numerical information more quickly than qualitative information since the numerical information may be comparatively more certain.

Finding significance for the conference call tone measures may indicate that we are simply picking up the impact of the earnings press releases, which is shown to be significantly related to abnormal returns by Davis et al. (2008), Demers and Vega (2008), Henry (2008), and Sadique (2008). To differentiate between the earnings announcements and the corresponding conference calls, we control for the tone of the press releases by separating the conference calls into two sections as in Matsumoto et al. (2007).[28] Kimbrough (2005) and Matsumoto et al. (2007) contend that the prepared statements by management at the beginning of a call essentially reiterate the information in the carefully crafted press release. Thus, we use the tone of the introductory statements prepared by management as a proxy for the tone of the earnings announcements. Following the prepared remarks, the remainder of the time is opened up and management fields questions from call participants. Accordingly, we incorporate new tone measures which represent these separate conference call components for both of the dictionaries used in this study (INTRO TONE and Q&A TONE, as well as INTRO H-TONE and Q&A H-TONE).

INTRO TONE is constructed exactly as the TONE variable above, only it is limited to the manager's introductory presentation text. Likewise, Q&A TONE represents the TONE measure for the question and answer portion of the call. INTRO H-TONE and Q&A H-TONE follow the same pattern using the Henry dictionary. In light of previous studies which demonstrate that conference calls convey information beyond that contained in the earnings announcements (Frankel et al., 1999; Kimbrough, 2005), we expect that, after controlling for the tone of the press releases using the conference call introductory statements, the tone of the question and answer dialogue between management and analysts will contain explanatory power for the cumulative abnormal returns. We expect this relation to be particularly pronounced for firms with greater investor cash flow uncertainty (non-dividend payers) who are more likely to rely on non-financial information than firms with less uncertainty (Amir and Lev, 1996).

We then conduct additional analysis at the portfolio level by forming tone quintiles based on the tone of the two call components separately. We examine mean and median CAR differences for the high and low quintile portfolios and compare the results for the INTRO and Q&A portions of the call.

[25] We also check our results using a consensus analyst forecast based surprise measure and obtain qualitatively similar results, albeit with considerable sample attrition.

[26] The additional variables are selected following related work by Tetlock (2007), Davis et al. (2008), Tetlock et al. (2008), Engelberg (2008), and Frankel et al. (2010), which control for both the disclosure of additional information and other factors that are known to affect returns.

[27] Firm effects denote dependence within observations of a firm across time. Time effects refer to dependence within a quarter across firms.

[28] Matsumoto et al. (2007) examine the relation between firm performance and quantity and type of disclosures made in conference calls. They split conference calls into their two primary components (the manager's introductory presentation and the question and answer session) and count the length of the calls, the number of financial versus non-financial words, and the number of backward-looking versus forward-looking verbs.

Table 3
Descriptive statistics and correlations.

	CAR(−1,1)	CAR(2,60)	SURP	$TONE_1$	$TONE_2$	H-$TONE_1$	H-$TONE_2$
Panel A							
Mean	0.05	−0.62	0.07	0.00	−0.00	5.21	0.64
Min	−51.35	−128.94	−83.48	−2.94	−4.95	0.82	−0.10
P25	−3.51	−8.65	−0.33	−0.94	−0.89	3.56	0.56
P50	0.12	−0.89	0.14	−0.18	0.04	4.73	0.65
P75	4.00	7.49	0.58	0.74	0.94	6.28	0.73
Max	48.64	89.92	89.47	15.16	5.55	28.00	1.00
Std.Dev.	7.54	14.98	5.00	1.36	1.36	2.43	0.13
N	2880	2880	2880	2880	2880	2880	2880
Panel B							
CAR(−1,1)	1.00						
CAR(2,60)	0.02	1.00					
SURP	0.10	−0.02	1.00				
$TONE_1$	0.09	0.01	0.04	1.00			
$TONE_2$	0.09	0.01	0.05	0.96	1.00		
H-$TONE_1$	0.10	0.04	0.06	0.59	0.56	1.00	
H-$TONE_2$	0.11	0.06	0.09	0.55	0.59	0.87	1.00

This table provides descriptive statistics (Panel A) and correlations (Panel B) for the full sample of 2880 conference calls. CAR(−1,1) is the 3-day cumulative abnormal return, in percent, where day 0 is the conference call date, where the abnormal returns are estimated using size-adjusted returns calculated as $AR_{j,t} = R_{j,t} - R_{p,t}$, where the abnormal return for firm j on day t is the difference between the return for firm j on day t and the mean return on day t for all firms in the same size decile as firm j. CAR(2,60) is calculated in the same manner as CAR(−1,1) only it is cumulated from days 2 through 60. SURP is the earnings surprise, in percent, calculated as {[EPS(qtr) − EPS(qtr-4)]/Stock Price (end of qtr-4) ∗ 100}. $TONE_1$ is the first principal component of the four ratios Positive/Negative, Strong/Weak, Active/Passive, and Overstated/Understated and $TONE_2$ is the first principal component of the four ratios (Positive − Negative)/(Positive + Negative), (Strong − Weak)/(Strong + Weak), (Active − Passive)/(Active + Passive), and (Overstated − Understated)/(Overstated + Understated), where each category reflects the proportion of words in a given call as defined by the Harvard IV-4 Psychosocial Dictionary. H-$TONE_1$ is the ratio of Positive words to Negative words and H-$TONE_2$ is the ratio (Positive − Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008).

While return is an important indicator of the level of asymmetry being dealt with, we confirm the effects call tone on actual trading using measures of cumulative abnormal volume (CAV). Following Barber and Odean (2008), we compute abnormal trading volume for firm j on day t, $AV_{j,t}$, as:

$$AV_{j,t} = \frac{V_{j,t}}{\overline{V}_{j,t}} \quad (15)$$

where $V_{j,t}$ is the volume for firm j on day t, and $\overline{V}_{j,t}$ is the mean volume for firm j from day $t = -252$ to $t = -2$ calculated as:

$$\overline{V}_{j,t} = \sum_{d=t-252}^{t-2} \frac{V_{j,d}}{252} \quad (16)$$

$AV_{j,t}$ is then cumulated over the 3 day initial reaction window, $CAV(-1,1)_j$, and days 2 through 60, $CAV(2,60)_j$, in the same manner as abnormal returns shown in Eqs. (11) and (12).[29]

We run firm level, cross-sectional regressions following Eq. (14), only we substitute CAV_j for CAR_j as the dependent variable. In its original form, a positive and significant tone variable in Eq. (14) indicates that returns increase as call tone increases. However, with CAV_j as the dependent variable, the interpretation is a little different since we expect increased abnormal trading volume when the sentiment reflected in the call language is either extremely positive or extremely negative. As such, significant tone variables in these regressions reveal the extent to which trading volume reacts asymmetrically to positive verses negative call tone.

5. Results

Descriptive statistics for the CARs, earnings surprise, and various tone measures are shown in Table 3, Panel A. The percent CARs for the initial reaction period range from −51.35 to 48.64, although most of the observations are less dramatic with a 25th percentile of −3.51 and a 75th percentile of 4.00. The drift period CARs, in percent, are more pronounced with a range of −128.94 to 89.92. The drift period interquartile range also implies that the majority of observations are less extreme with the 25th percentile at −8.65 and the 75th percentile at 7.49.[30] However, regardless of whether we look at the full range or interquartile range, CAR(2,60) is roughly double the levels we see with CAR(−1,1), suggesting that the information content of the earnings surprises and accompanying conference calls are not fully impounded into prices during the initial reaction period.

The mean (median) percent SURP of .07 (.14) is near zero and ranges from -83.48 to 89.47 which indicates that the sample is fairly evenly distributed between positive and negative earnings surprises, albeit with a slight negative skew. The tone measures reveal that the overall tone of the conference calls in the sample is positive. With a mean of 5.21, H-$TONE_1$ shows that the conference call participants use over five times more positive words than negative words, on average. In the extreme positive case, call tone is 28 times more positive than negative, while in the extreme negative case a ratio of .82 indicates that call tone is only slightly more negative than positive. This is not surprising since we would expect management to put their best foot forward and present their case in the best possible light. However, the overall positivity of the sample is not nearly as important as the substantial dispersion we see in the tone measures.[31] The H-$TONE_2$ measure, which is bounded between −1 and 1, shows a minimum value of −.10 and a maximum value of 1. The minimum value indicates a call tone that is slightly more negative than positive, but is near the point of tone neutrality. However, there is only one observation out of nearly three thousand in our sample that has an H-$TONE_2$ measure less than zero. Likewise, there is only one call with an H-$TONE_2$ measure of 1.00, which represents a call that is completely void of negative words. Although not shown, the 1st and 99th percentiles are at .27 and .86, respectively. Descriptive statistics for $TONE_1$ and $TONE_2$ show the same levels of dispersion even though the numerical rep-

[29] Results are not sensitive to the method with which we construct abnormal volume.

[30] Later results are robust to winsorising the extreme CARs at both 1% and 5%.

[31] We could also limit our focus to only the proportion of negative words as in Tetlock (2007), Tetlock et al. (2008), and others. However, we contend that it is not so important how positive or negative a call is in and of itself. Rather, the relative proportions of the two tonal dimensions taken together provide a richer information set for our analysis.

Table 4
Test of differences of means and medians, by $TONE_1$ and $TONE_2$ quintiles.

$TONE_1$ quintiles		CAR(−1, 1)	CAR(2, 60)	$TONE_2$ quintiles	CAR(−1, 1)	CAR(2, 60)
1 (Low)	Mean	−1.04	−0.90	1 (Low)	−1.16	−1.38
	Median	−1.00	−1.44		−1.07	−1.56
2	Mean	−0.28	−0.54	2	0.00	0.45
	Median	0.08	−0.06		0.29	0.63
3	Mean	0.51	−0.93	3	0.05	−1.44
	Median	0.51	−0.98		0.08	−1.46
4	Mean	0.27	−1.18	4	0.65	−0.78
	Median	0.20	−1.54		0.36	−1.24
5 (High)	Mean	0.79	0.45	5 (High)	0.71	0.06
	Median	0.88	−0.01		0.89	−0.59
Mean Q5–Q1	1.83***	1.35		1.87***	1.43	
t-Statistic	(4.16)	(1.59)		(4.24)	(1.62)	
Wilcoxon rank-sum test						
Median Q5–Q1	1.88***	1.43*		1.97***	0.97	
z-Statistic	(5.40)	(1.79)		(5.86)	(1.48)	

This table shows the differences in CARs when sorted into $TONE_1$ and $TONE_2$ quintile portfolios. Test statistics (*t* and *z*) are in parentheses. CAR(−1, 1) is the 3-day cumulative abnormal return, in percent, where day 0 is the conference call date, where the abnormal returns are estimated using size-adjusted returns calculated as $AR_{j,t} = R_{j,t} - R_{p,t}$, where the abnormal return for firm *j* on day *t* is the difference between the return for firm *j* on day *t* and the mean return on day *t* for all firms in the same size decile as firm *j*. CAR(2, 60) is calculated in the same manner as CAR(−1, 1) only it is cumulated from days 2 through 60. $TONE_1$ is the first principal component of the four ratios Positive/Negative, Strong/Weak, Active/Passive, and Overstated/Understated, where each category reflects the proportion of words in a given call as defined by the Harvard IV-4 Psychosocial Dictionary. $TONE_2$ is the first principal component of the four ratios (Positive − Negative)/(Positive + Negative), (Strong − Weak)/(Strong + Weak), (Active − Passive)/(Active + Passive), and (Overstated − Understated)/(Overstated + Understated), where each category reflects the proportion of words in a given call as defined by the Harvard IV-4 Psychosocial Dictionary.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

resentations are somewhat different because we use the first principal component of the four ratios.

The correlations in Panel B of Table 3 show that the tone measures are all highly correlated with one another. However, the measures obtained using the Harvard and Henry dictionaries are not perfectly correlated, providing an indication that the two dictionaries materially differ from one another. The two dictionaries also differ with respect to the level of correlation between the tone measures and both the initial reaction window CARs and drift window CARs. With coefficients of .10 and .11, H-$TONE_1$ and H-$TONE_2$ are slightly more correlated with CAR(−1, 1) than the Harvard dictionary based $TONE_1$ (.09) and $TONE_2$ (.09) measures. The difference is more distinct when focusing on the drift period. The correlation coefficients for H-$TONE_1$ and H-$TONE_2$ with CAR(2, 60) are .04 and .06, respectively. While the coefficients for $TONE_1$ and $TONE_2$ with CAR(2, 60) are both .01.

Differences between the two dictionaries are further seen in the preliminary analysis of Tables 3 and 4 which show CARs sorted into $TONE_1$ and $TONE_2$ (Table 4) and H-$TONE_1$ and H-$TONE_2$ (Table 5) quintile portfolios. The Harvard dictionary based tone measures in Table 4 produce highly significant high-minus-low CAR differences of just under 2% during the initial reaction window, but lose most of their significance and some of their magnitude during the drift period. In contrast, we find the Henry dictionary measures in Table 5 to be much more powerful in their ability to detect differences in the context of earnings conference calls. With the finance-specific dictionary, we find highly significant CARs during both the initial reaction window and the drift period, each in excess of 2%. This provides out-of-sample support for the argument of Loughran and McDonald (2011) that customized dictionaries with context specific terminology are more appropriate for research in financial settings. Accordingly, we focus on the results obtained using the earnings specific Henry dictionary from this point forward. Specifically, we show results using the H-$TONE_2$ measure, which are either identical or nearly identical to the results obtained using the H-$TONE_1$ measure. With H-$TONE_1$ and H-$TONE_2$, the sorting breakpoints are identical for the highest and lowest quintile portfolios. The middle quintile portfolio breakpoints are so close between the two H-TONE measures that only minor differences can be seen. This results in identical zero investment strategy returns in terms of both magnitude and significance across the two H-TONE measures.[32]

Fig. 1 provides a graphical illustration of the CAR differences between the high and low H-$TONE_2$ quintile portfolios plotted against the average of all conference calls in the sample. We see the initial reaction followed by continued spreading of the two portfolios through 60 days after the calls. The total spread of roughly 5.5% at the end of the drift period is dramatically different than the overall average which hovers near zero.

In Fig. 2 we demonstrate performance variation based on differences in earnings surprises. Similar to the graphical representation in Fig. 1, we plot the cumulative abnormal returns from 10 days prior to the conference call through 60 days after the conference call for firms in the high SURP quintile, firms in the low SURP quintile, and an average of all firms in the sample. With a total spread of approximately 4.5% at the end of the drift period, Fig. 2 illustrates CAR differences in the cross section of earnings surprises of a magnitude similar to, but slightly less than, we see in the H-$TONE_2$ cross section.

Fig. 3 illustrates the differences between H-$TONE_2$ portfolios while controlling for SURP. The CARs are separated into four separate portfolios corresponding to high and low H-$TONE_2$ terciles and high and low SURP terciles. When holding SURP constant we see that there is a remarkable difference between CARs based on the level of H-$TONE_2$. That is, when the earnings surprises fall within the lowest third of all sample surprises, firms with low call tone see abnormal returns drop precipitously while those firms with high call tone are largely unaffected by the negative surprise. Essentially, a positive call tone appears to offset the damaging

[32] We choose H-$TONE_2$ because it is bounded between −1 and 1 and is normally distributed, while H-$TONE_1$ is unbounded and leptokurtic. Results obtained using the Harvard dictionary based $TONE_1$ and $TONE_2$ measures are less powerful, but qualitatively the same.

Table 5
Test of differences of means and medians, by H-TONE$_1$ and H-TONE$_2$ quintiles.

H-TONE$_1$ quintiles		CAR(−1,1)	CAR(2,60)	H-TONE$_2$ quintiles	CAR(−1,1)	CAR(2,60)
1 (Low)	Mean	−1.25	−1.86	1 (Low)	−1.25	−1.86
	Median	−0.95	−1.46		−0.95	−1.46
2	Mean	−0.14	−1.13	2	−0.17	−1.11
	Median	−0.45	−1.61		−0.47	−1.61
3	Mean	−0.08	−0.04	3	−0.06	−0.06
	Median	0.15	−0.72		0.16	−0.74
4	Mean	0.54	−0.42	4	0.54	−0.42
	Median	0.41	−0.74		0.41	−0.74
5 (High)	Mean	1.19	0.35	5 (High)	1.19	0.35
	Median	1.37	0.56		1.37	0.56
Mean Q5–Q1		2.44***	2.21**		2.44***	2.21**
t-Statistic		(5.43)	(2.57)		(5.43)	(2.57)
Wilcoxon rank-sum test						
Median Q5–Q1		2.32***	2.01***		2.32***	2.01***
z-Statistic		(6.53)	(2.85)		(6.53)	(2.85)

This table shows the differences in CARs when sorted into H-TONE$_1$ and H-TONE$_2$ quintile portfolios. Test statistics (t and z) are in parentheses. CAR(−1,1) is the 3-day cumulative abnormal return, in percent, where day 0 is the conference call date, where the abnormal returns are estimated using size-adjusted returns calculated as $AR_{j,t} = R_{j,t} - R_{p,t}$, where the abnormal return for firm j on day t is the difference between the return for firm j on day t and the mean return on day t for all firms in the same size decile as firm j. CAR(2,60) is calculated in the same manner as CAR(−1,1) only it is cumulated from days 2 through 60. H-TONE$_1$ is the ratio of positive words to negative words as defined by the custom earnings dictionary of Henry (2008). H-TONE$_2$ is the ratio of (Positive − Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008).

*** $p < 0.01$.

** $p < 0.05$.

* $p < 0.1$.

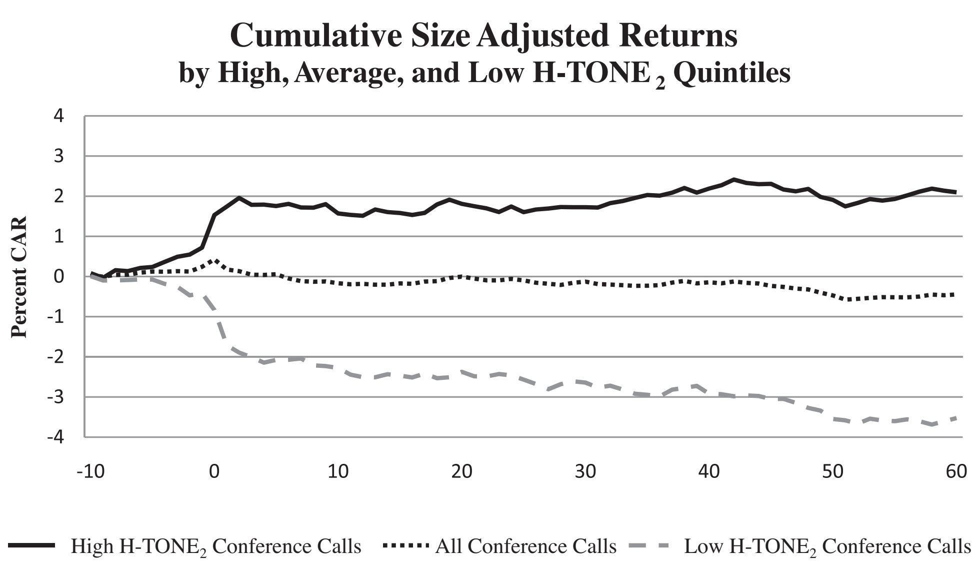

Fig. 1. Cumulative abnormal returns (CARs) across all conference calls and by high and low H-TONE$_2$ quintiles. Abnormal returns are estimated using returns that have been size-adjusted and are calculated as $\{AR_{j,t} = R_{j,t} - R_{pt}\}$ where $AR_{j,t}$ is the abnormal return for firm j, day t; $R_{j,t}$ is the return for firm j, day t; and R_{pt} is the equally weighted mean return on day t for all firms in the same size decile as firm j. Abnormal returns are cumulated over a 71 day period starting 10 days prior to the earnings conference call through 60 days after the call. Day zero is the date of the earnings conference call. H-TONE$_2$ is the ratio of (Positive − Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008). High H-TONE$_2$ Conference Calls shows the CARs for the highest H-TONE$_2$ quintile conference calls. Low H-TONE$_2$ Conference Calls shows the CARs for the lowest H-TONE$_2$ quintile conference calls. All conference calls represents the CARs for all conference calls in the sample. CARs are shown in percent.

effect of a poor earnings announcement. When the earnings surprises fall within the highest third of all sample surprises, firms with high call tone experience strong cumulative abnormal returns. However, those firms with high surprises and low call tone see returns that increase very little. By comparison, the CARs are indistinguishable from the High H-TONE$_2$ Low SURP portfolio by day $t = 60$. In contrast, the total difference between the High H-TONE$_2$ High SURP portfolio and the Low H-TONE$_2$ Low SURP is roughly 8% on a quarterly basis.

At the firm level, the separate effects of tone and earnings surprises for both the initial reaction period CARs and the drift period CARs are shown in the regression results of Table 6. The results for all firms in the sample are in Panel A. As expected, the coefficients for both SURP and H-TONE$_2$ are positive and highly significant

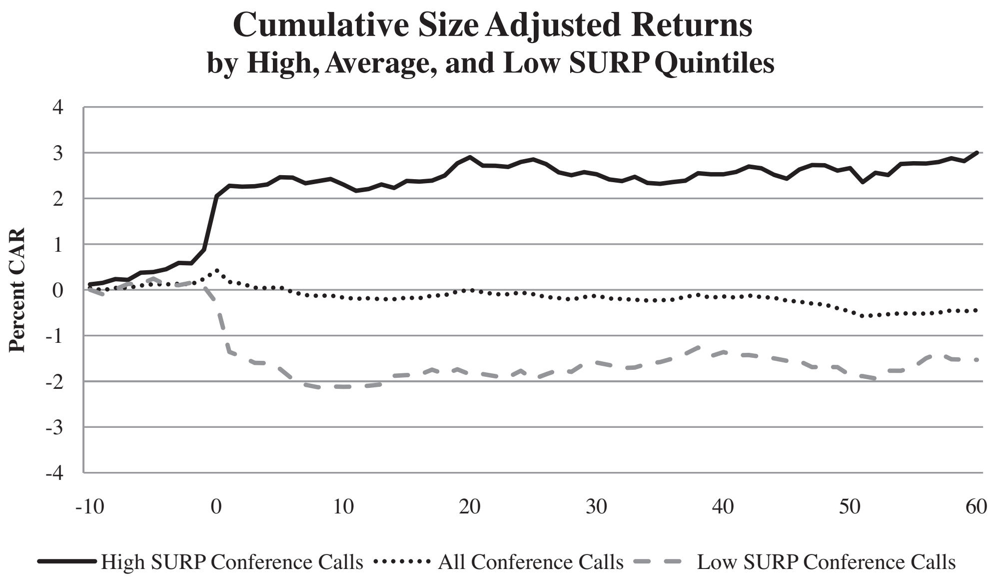

Fig. 2. Cumulative abnormal returns (CARs) across all conference calls and by high and low SURP quintiles. Abnormal returns are estimated using returns that have been size-adjusted and are calculated as $\{AR_{j,t} = R_{j,t} - R_{pt}\}$ where $AR_{j,t}$ is the abnormal return for firm j, day t; $R_{j,t}$ is the return for firm j, day t; and R_{pt} is the equally weighted mean return on day t for all firms in the same size decile as firm j. Abnormal returns are cumulated over a 71 day period starting 10 days prior to the earnings conference call through 60 days after the call. Day zero is the date of the earnings conference call. SURP is the earnings surprise calculated as {[EPS(qtr) – EPS(qtr-4)]/Stock Price (end of qtr-4)}. High SURP Conference Calls shows the CARs for the highest SURP quintile conference calls. Low SURP Conference Calls shows the CARs for the lowest SURP quintile conference calls. All Conference Calls represents the CARs for all conference calls in the sample. CARs are shown in percent.

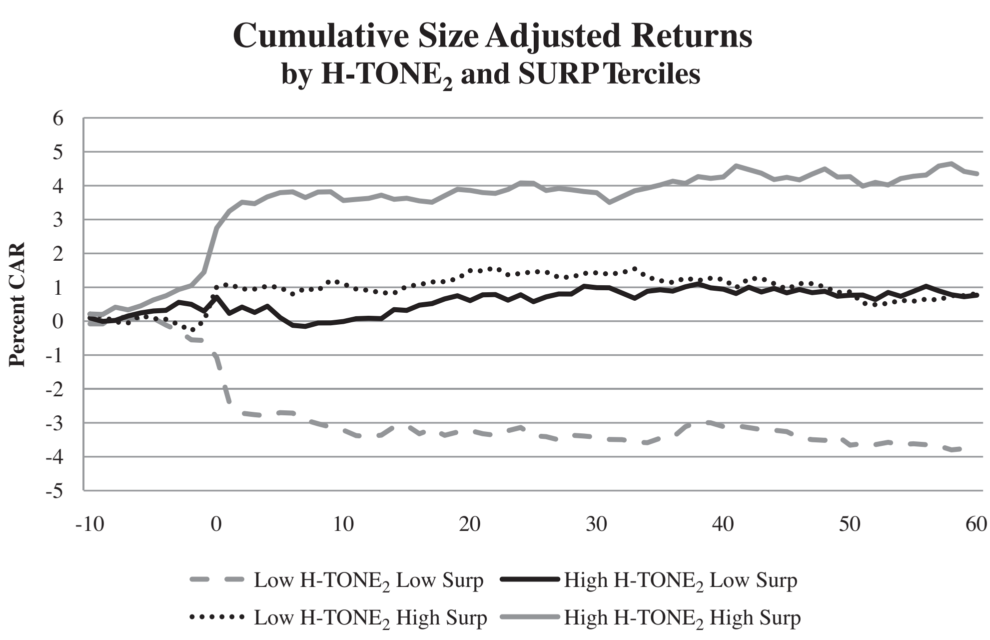

Fig. 3. Cumulative abnormal returns by earnings surprise and H-TONE$_2$ terciles. Abnormal returns are estimated using returns that have been size-adjusted and are calculated as $\{AR_{j,t} = R_{j,t} - R_{pt}\}$ where $AR_{j,t}$ is the abnormal return for firm j, day t; $R_{j,t}$ is the return for firm j, day t; and R_{pt} is the equally weighted mean return on day t for all firms in the same size decile as firm j. Abnormal returns are cumulated over a 71 day period starting 10 days prior to the earnings conference call through 60 days after the call. Day zero is the date of the earnings conference call. Cumulative abnormal returns (CAR) are shown in percent. H-TONE$_2$ is the ratio of (Positive – Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008). SURP is the earnings surprise calculated as {[EPS(qtr) – EPS(qtr-4)]/Stock Price (end of qtr-4)}. CARs are plotted for High H-TONE$_2$ High SURP conference calls, High H-TONE$_2$ Low SURP conference calls, Low H-TONE$_2$ High SURP conference calls, and Low H-TONE$_2$ Low SURP conference calls. High H-TONE$_2$ High SURP contains the conference calls within the highest H-TONE$_2$ tercile and lowest SURP tercile. High H-TONE$_2$ Low SURP contains the conference calls within the highest H-TONE$_2$ tercile and lowest SURP tercile. Low H-TONE$_2$ High SURP contains the conference calls within the lowest H-TONE$_2$ tercile and highest SURP tercile. Low H-TONE$_2$ Low SURP contains the conference calls within the lowest H-TONE$_2$ tercile and lowest SURP tercile.

when the dependent variable is CAR(−1,1). When the dependent variable is CAR(2,60) tone remains an important predictor, however SURP loses its significance. Consistent with Engelberg (2008) and Demers and Vega (2008), qualitative information appears to have greater explanatory power for returns at longer horizons than the actual earnings number. In other words, numerical data may be easier for markets to comprehend. When conference calls are separated into their two main components, we find that significant

Table 6
Regression results of cumulative abnormal returns on earnings surprise and H-TONE$_2$.

	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(2,60)	CAR(2,60)	CAR(2,60)	CAR(2,60)	CAR(2,60)	CAR(2,60)
Panel A: All												
SURP	0.148***	0.133***	0.135***	0.145***	0.137***	0.103**	−0.047	−0.064	−0.065	−0.051	−0.062	−0.132
	(5.13)	(4.10)	(4.21)	(4.73)	(4.21)	(2.28)	(−0.45)	(−0.64)	(−0.66)	(−0.50)	(−0.62)	(−1.28)
H-TONE$_2$		6.213***						6.942**				
		(5.37)						(2.20)				
INTRO H-TONE$_2$			3.412***		2.343***	1.572*			4.884***		3.218*	3.989**
			(4.41)		(2.90)	(1.77)			(2.67)		(1.77)	(2.21)
Q&A H-TONE$_2$				4.362***	3.202***	3.478***				6.581**	4.988*	4.666*
				(3.89)	(2.68)	(2.93)				(2.54)	(1.89)	(1.92)
INTRO COUNT						−0.040						0.356
						(−0.34)						(1.41)
Q&A COUNT						−0.073						−0.370***
						(−0.99)						(−2.73)
SIZE						0.365*						−0.708
						(1.67)						(−1.32)
BM						0.045						−0.038
						(0.76)						(−0.87)
ROA						0.115*						0.200*
						(1.73)						(1.66)
LEVERAGE						0.022**						0.006
						(2.02)						(0.29)
VOLUME						−0.274						−0.151
						(−1.45)						(−0.37)
VOLATILITY						−0.145						−0.515
						(−0.62)						(−0.63)
ANALYST						0.002						−0.678**
						(0.02)						(−2.16)
DECLARATION						0.628*						0.473
						(1.86)						(0.60)
Constant	0.040	−3.904***	−2.157***	−2.689***	−3.471***	−2.766	−0.617	−5.025***	−3.761***	−4.735***	−5.810***	3.565
	(0.25)	(−5.32)	(−4.31)	(−3.82)	(−4.70)	(−1.42)	(−1.46)	(−2.63)	(−3.24)	(−3.16)	(−3.49)	(0.74)
Observations	2880	2880	2880	2880	2880	2879	2880	2880	2880	2880	2880	2879
R-Squared	0.010	0.020	0.015	0.016	0.018	0.033	0.000	0.004	0.003	0.004	0.005	0.024
Panel B: Dividend												
SURP	0.111**	0.071*	0.076*	0.092**	0.072*	−0.015	−0.027	−0.068	−0.081	−0.041	−0.081	−0.123
	(2.27)	(1.74)	(1.74)	(2.01)	(1.68)	(−0.19)	(−0.15)	(−0.38)	(−0.46)	(−0.23)	(−0.47)	(−0.48)
H-TONE$_2$		8.605***						8.696***				
		(5.82)						(2.79)				
INTRO H-TONE$_2$			4.817***		3.600***	3.541***			7.423***		7.213***	7.834***
			(5.15)		(3.96)	(3.33)			(3.92)		(3.72)	(4.60)
Q&A H-TONE$_2$				5.788***	3.886***	4.031***				4.485	0.673	0.442
				(4.33)	(2.84)	(2.94)				(1.57)	(0.22)	(0.14)
INTRO COUNT						0.111						−0.224
						(0.89)						(−0.73)
Q&A COUNT						−0.072						−0.228
						(−0.90)						(−1.03)
SIZE						0.132						0.411
						(0.48)						(0.59)
BM						1.457**						0.894
						(2.28)						(0.47)
ROA						0.206						0.085
						(1.47)						(0.36)
LEVERAGE						0.033***						−0.004
						(2.66)						(−0.14)
VOLUME						−0.190						−0.281
						(−0.77)						(−0.69)
VOLATILITY						0.267						0.564
						(0.70)						(0.51)
ANALYST						0.259						−1.268**
						(1.57)						(−2.54)
DECLARATION						0.700*						0.271
						(1.71)						(0.32)
Constant	0.103	−5.273***	−2.957***	−3.459***	−4.575***	−6.743***	−0.900**	−6.332***	−5.616***	−3.660*	−5.896***	−3.284
	(0.45)	(−5.56)	(−4.61)	(−3.95)	(−4.67)	(−2.82)	(−2.39)	(−2.91)	(−3.87)	(−1.87)	(−2.77)	(−0.57)
Observations	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440
R-Squared	0.006	0.033	0.022	0.020	0.028	0.043	0.000	0.008	0.011	0.002	0.011	0.029
Panel C: NonDiv												
SURP	0.174***	0.171***	0.172***	0.179***	0.177***	0.157***	−0.065	−0.068	−0.067	−0.054	−0.052	−0.117
	(4.17)	(4.06)	(4.10)	(4.22)	(4.23)	(3.79)	(−0.47)	(−0.49)	(−0.48)	(−0.39)	(−0.37)	(−0.91)
H-TONE$_2$		4.107***						4.726				
		(2.61)						(1.18)				

Table 6 (*continued*)

	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(2,60)	CAR(2,60)	CAR(2,60)	CAR(2,60)	CAR(2,60)	CAR(2,60)
INTRO H-TONE$_2$			2.040*		1.023	−0.232			1.651		−1.434	−1.448
			(1.88)		(0.75)	(−0.15)			(0.56)		(−0.43)	(−0.39)
Q&A H-TONE$_2$				3.397**	2.937	3.462*				8.261***	8.906***	9.185***
				(2.12)	(1.53)	(1.78)				(2.76)	(2.61)	(2.95)
INTRO COUNT						−0.214						0.860***
						(−1.01)						(2.84)
Q&A COUNT						−0.107						-0.623**
						(−1.05)						(−2.42)
SIZE						0.742***						−1.655**
						(2.77)						(−2.55)
BM						0.027						−0.078**
						(0.65)						(−2.56)
ROA						0.119**						0.292**
						(2.16)						(1.98)
LEVERAGE						0.020						0.013
						(1.49)						(0.64)
VOLUME						−0.435*						−0.075
						(−1.95)						(−0.14)
VOLATILITY						−0.286						−1.45*
						(−0.90)						(−1.67)
ANALYST						−0.209						−0.098
						(−0.90)						(−0.23)
Constant	−0.033	−2.685***	−1.366*	−2.196**	−2.572***	−0.471	−0.328	−3.380	−1.407	−5.588***	−5.062***	11.050*
	(−0.12)	(−2.67)	(−1.88)	(−2.08)	(−2.63)	(−0.20)	(−0.40)	(−1.47)	(−0.85)	(−3.39)	(−2.81)	(1.68)
Observations	1440	1440	1440	1440	1440	1439	1440	1440	1440	1440	1440	1439
R-Squared	0.013	0.016	0.014	0.016	0.016	0.041	0.000	0.002	0.001	0.005	0.005	0.039

This table provides results for cross-sectional regressions of CARs on SURP and measures of H-TONE$_2$ for the full sample (Panel A), dividend paying firms (Panel B), and firms that do not pay dividends (Panel C). CAR(−1,1) is the 3-day cumulative abnormal return, in percent, where day 0 is the conference call date, where the abnormal returns are estimated using size-adjusted returns calculated as $AR_{j,t} = R_{j,t} - R_{p,t}$, where the abnormal return for firm *j* on day *t* is the difference between the return for firm *j* on day *t* and the mean return on day *t* for all firms in the same size decile as firm *j*. CAR(2,60) is calculated in the same manner as CAR(−1,1) only it is cumulated from days 2 through 60. SURP is the earnings surprise, in percent, calculated as {[EPS(qtr) − EPS(qtr-4)]/Stock Price (end of qtr-4) ∗ 100}. H-TONE$_2$ is the ratio of (Positive − Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008). The variable prefixes INTRO and Q&A indicate that the referenced tone measures are for the introductory managerial statement and the question and answer portions of the conference calls, respectively. COUNT is the number of words, in thousands, for a given portion of a conference call (INTRO and Q&A). SIZE is the log of firm market capitalization from the previous quarter. BM is the ratio of book-to-market equity as of the end of the previous quarter. ROA is return on assets, in percent, calculated as net income divided by total assets multiplied by one hundred. LEVERAGE is expressed in percent as the ratio of total liabilities to total assets multiplied by one hundred. VOLUME is the log of the total share trading volume on day zero. VOLATILITY is in percent and is calculated as the standard deviation of daily returns for the ninety trading-day period ending 10 days prior to the conference call multiplied by one hundred. ANALYST is the log of the number of analysts who cover a given firm. DECLARATION is an indicator variable equal to one if the firm declares dividends within the 3-day window around the conference call (−1,1), and zero otherwise. Standard errors are adjusted for heteroscedasticity following White (1980) and for clustering by firm and quarter following Petersen (2009). *t*-Statistics are in parentheses.

*** $p < 0.01$.

** $p < 0.05$.

* $p < 0.1$.

additional information is contained in the question and answer portion of the calls after controlling for the prepared managerial remarks. After controlling for the linguistic information in the earnings announcement, the tone of the conference calls do indeed provide additional value relevant information to market participants. These results hold after including additional variables known to affect returns and control for information disclosure for both the initial reaction window and the drift period.[33] Thus, using content analysis techniques we find support for the arguments of Frankel et al. (1999) and Kimbrough (2005).[34]

Table 6 also provides regression results for dividend paying firms (Panel B) and firms that do not pay dividends (Panel C). In both cases, tone appears to have significant explanatory power, though there are important differences between the two when separating the calls into their two components. With dividend payers, we find that the tone in the introductory portion of the conference calls (INTRO H-TONE$_2$) has direct explanatory power during both the initial reaction and drift windows. However, while Q&A H-TONE$_2$ matters initially, it loses its significance during the CAR(2,60) period. In contrast, Q&A H-TONE$_2$ is highly significant during the latter window for firms that do not pay dividends, suggesting that the conference call question and answer discussion is relied on more heavily in cases of greater cash flow uncertainty.

The portfolio level analysis in Table 7 underscores the advantage of call component separation and confirms the differences in the importance investors place on the question and answer session tone for dividend paying and non-paying firms. Panel A shows that while INTRO H-TONE$_2$ conveys valuable information, mean portfolio differences are not significant during the drift window for firms that do not pay dividends. The opposite result is seen in Panel B where Q&A H-TONE$_2$ based portfolio sorts reveal dividend payers to have an insignificant mean CAR(2,60) portfolio difference of 1.68%, compared to a significant mean CAR(2,60) portfolio difference of 2.94%. Thus separating conference calls into their two primary components provides us with a more accurate understanding of the relation between conference call tone and stock returns relative to investor cash flow uncertainty.

Table 8 provides results from regressing cumulative abnormal volume on call characteristics, such as tone and word count, and other control variables. In all three panels (A–C) Q&A COUNT, which measures the length of the question and answer portion of the call, is positive and highly significant. Given that longer calls result in more initial trading, we can infer that conference call disclosures convey value relevant information, consistent with prior

[33] Although not shown, we repeat the analysis and exclude all firms who both declare dividends within the initial reaction window (−1,1) and whose dividend changes from the prior quarter. The results are not materially affected.

[34] Frankel et al. (1999) and Kimbrough (2005) do not use content analysis or examine conference call characteristics. Rather they examine return volatility and trading volume at the time of the call (Frankel et al.) and the initiation of conference calls (Kimbrough).

Table 7
Test of differences of means and medians, by H-TONE$_2$ component quintiles.

Quintile		CAR(−1,1)	CAR(−1,1)	CAR(−1,1)	CAR(2,60)	CAR(2,60)	CAR(2,60)
		All	Dividend	NonDiv	All	Dividend	NonDiv
Panel A: INTRO H-TONE$_2$ portfolios							
1 (Low)	Mean	−1.18	−1.88	−0.42	−2.13	−2.78	−1.32
	Median	−1.11	−1.59	−0.02	−2.02	−2.12	−1.62
5 (High)	Mean	1.08	1.22	0.87	0.48	0.66	0.03
	Median	1.41	1.25	1.78	0.04	0.56	−0.70
Mean Q5–Q1		2.26***	3.10***	1.29*	2.61***	3.44***	1.35
t-Statistic		(5.14)	(5.92)	(1.83)	(2.93)	(3.10)	(0.97)
Wilcoxon rank-sum test							
Median Q5–Q1		2.52***	2.84***	1.80**	2.07***	2.68***	0.91
z-Statistic		(6.26)	(6.40)	(2.56)	(2.86)	(2.93)	(0.85)
Panel B: Q&A H-TONE$_2$ portfolios							
1 (Low)	Mean	−0.83	−0.92	−0.77	−1.48	−1.23	−1.93
	Median	−0.64	−0.69	−0.48	−1.38	−1.08	−1.63
5 (High)	Mean	1.12	1.38	0.99	0.58	0.44	1.02
	Median	1.21	0.86	1.75	0.74	0.41	1.52
Mean Q5–Q1		1.95***	2.30***	1.76**	2.06**	1.68	2.94**
t-Statistic		(4.37)	(4.13)	(2.55)	(2.29)	(1.61)	(2.02)
Wilcoxon rank-sum test							
Median Q5–Q1		1.84***	1.55***	2.23***	2.11***	1.49*	3.16**
z-Statistic		(5.01)	(3.97)	(3.17)	(2.60)	(1.78)	(2.01)

This table shows the differences in CARs when sorted into INTRO H-TONE$_2$ (Panel A) and Q&A H-TONE$_2$ (Panel B) quintile portfolios for the full sample (All), dividend paying firms (Dividend), and firms that do not pay dividends (NonDiv). Test statistics (t and z) are in parentheses. CAR(−1,1) is the 3-day cumulative abnormal return, in percent, where day 0 is the conference call date, where the abnormal returns are estimated using size-adjusted returns calculated as $AR_{j,t} = R_{j,t} - R_{p,t}$, where the abnormal return for firm j on day t is the difference between the return for firm j on day t and the mean return on day t for all firms in the same size decile as firm j. CAR(2,60) is calculated in the same manner as CAR(−1,1) only it is cumulated from days 2 through 60. H-TONE$_2$ is the ratio of (Positive − Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008).
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

studies. We further see that over the drift window call tone is positive and highly significant for the whole sample (Panel A), but that this is primarily driven by the subsample of firms with greater cash flow uncertainty (Panel C). The positive and significant tone variables over the longer horizon indicate that the market more slowly incorporates the information conveyed by positive call tone. When separated into INTRO and Q&A components, the Q&A H-TONE$_2$ dominates.

6. Conclusion

It is generally assumed that managers possess superior information regarding a firm's future prospects than investors (Healy and Palepu, 2001). The means through which such information becomes known and the extent and efficiency of the corresponding investor reaction are not fully understood (Core, 2001; Dye, 2001; Healy and Palepu, 2001; Verrecchia, 2001). This study extends the empirical disclosure literature by examining unique aspects of quarterly earnings conference calls and the subsequent market reaction. We study investor response in the context of uncertainty using computer aided textual analysis and the hand-collected transcripts of quarterly earnings conference calls.

We employ a general word categorization dictionary that is widely used across disciplines (Harvard IV-4 Psychosocial) and a custom, earnings specific dictionary (Henry) in order to quantify the linguistic tone (i.e. sentiment) of earnings conference call wording. We find that call tone is significantly related to the initial earnings announcement window abnormal stock returns, the post-earnings-announcement drift, and abnormal trading volume, after controlling for the numerical representation of the earnings surprise. However, the custom, "domain relevant" dictionary is shown to be much more powerful in detecting cumulative abnormal returns beyond the initial reaction window, providing out-of-sample support for the argument of Loughran and McDonald (2011) that context specific dictionaries are most appropriate for use in financial research.

Unlike previous studies, which either limit their focus to media expressed sentiment or management expressed tone, we examine the marginal contribution of the tone of the dialogue between management and analysts in helping us better understand the behavior of stock returns following earnings announcements. We differentiate between the two major components of an earnings conference call by separating management's prepared introductory remarks and the more spontaneous question and answer portion of the call. Since the prepared statements essentially reiterate the content of the earnings press release (Frankel et al., 1999; Kimbrough, 2005), including the two separate call components in our analysis allows us to determine the contribution of the informal, unscripted discussion between management and analysts above and beyond the information in the earnings number and accompanying announcement.

We find that the textual tone of the question and answer portion of the call has significant predictive ability for the initial reaction CARs, the post-earnings-announcement drift, and abnormal trading volume after controlling for both the numerical earnings surprise and the tone of the prepared remarks by management. Both firm-level and portfolio-level results suggest that the incremental information conveyed by conference call tone is relied upon more heavily by investors for firms that do not pay dividends. In other words, conference call question and answer tone has more explanatory power for abnormal returns and trading volume during the 60 trading days after the call when there is greater investor cash flow uncertainty.

Overall, our results point to several interesting areas for future research, which can help address the limitations of this study. The

Table 8

Regression results of cumulative abnormal volume on earnings surprise and H-TONE$_2$. This table provides results for cross-sectional regressions of CAV on SURP and measures of H-TONE$_2$ for the full sample (Panel A), dividend paying firms (Panel B), and firms that do not pay dividends (Panel C). CAV(−1,1) is the 3-day cumulative abnormal volume, where day 0 is the conference call date, and abnormal volume is calculated as $AV_{j,t} = V_{j,t}/\overline{V}_{j,t}$, where the abnormal volume for firm j on day t is the volume for firm j on day t divided by the mean volume for firm j from day $t = -252$ to $t = -2$. CAV(2,60) is calculated in the same manner as CAV(−1,1) only it is cumulated from days 2 through 60. SURP is the earnings surprise, in percent, calculated as {[EPS(qtr) − EPS(qtr-4)]/Stock Price (end of qtr-4) $*$ 100}. H-TONE$_2$ is the ratio of (Positive − Negative)/(Positive + Negative) where each category reflects the proportion of words in a given call as defined by the custom earnings dictionary of Henry (2008). The variable prefixes INTRO and Q&A indicate that the referenced tone measures are for the introductory managerial statement and the question and answer portions of the conference calls, respectively. COUNT is the number of words, in thousands, for a given portion of a conference call (INTRO and Q&A). SIZE is the log of firm market capitalization from the previous quarter. BM is the ratio of book-to-market equity as of the end of the previous quarter. ROA is return on assets, in percent, calculated as net income divided by total assets multiplied by one hundred. LEVERAGE is expressed in percent as the ratio of total liabilities to total assets multiplied by one hundred. VOLATILITY is in percent and is calculated as the standard deviation of daily returns for the ninety trading-day period ending 10 days prior to the conference call multiplied by one hundred. ANALYST is the log of the number of analysts who cover a given firm. DECLARATION is an indicator variable equal to one if the firm declares dividends within the 3-day window around the conference call (−1,1), and zero otherwise. Standard errors are adjusted for heteroscedasticity following White (1980) and for clustering by firm and quarter following Petersen (2009). t-Statistics are in parentheses.

	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (2,60)	CAV (2,60)	CAV (2,60)	CAV (2,60)	CAV (2,60)	CAV (2,60)
Panel A: All												
SURP	0.013	0.012	0.012	0.013	0.011	−0.014	0.103	0.060	0.062	0.093	0.071	−0.065
	(0.48)	(0.44)	(0.40)	(0.48)	(0.39)	(−0.35)	(0.62)	(0.33)	(0.35)	(0.53)	(0.40)	(−0.30)
H-TONE$_2$		0.163						18.145***				
		(0.20)						(3.13)				
INTRO H-TONE$_2$			0.347		0.456	0.839			11.232**		6.547	7.967
			(0.56)		(0.66)	(1.19)			(2.33)		(1.23)	(1.46)
Q&A H-TONE$_2$				−0.102	−0.328	−0.353				17.271***	14.030***	14.121***
				(−0.16)	(−0.44)	(−0.52)				(4.03)	(2.92)	(3.08)
INTRO COUNT						0.081						−0.648
						(1.18)						(−1.03)
Q&A COUNT						0.274***						−0.399
						(5.04)						(−1.04)
SIZE						−0.428***						−0.557
						(−4.91)						(−1.15)
BM						−0.042						−0.080
						(−0.98)						(−0.37)
ROA						0.075**						0.387
						(2.01)						(1.49)
LEVERAGE						−0.008*						0.093***
						(−1.90)						(3.59)
VOLATILITY						0.133						1.114
						(1.13)						(0.94)
ANALYST						0.006						−1.615*
						(0.07)						(−1.72)
DECLARATION						0.263						6.329**
						(0.85)						(2.02)
Constant	6.136***	6.032***	5.913***	6.200***	6.047***	7.545***	70.059***	58.539***	62.828***	59.253***	57.066***	58.911***
	(48.45)	(12.35)	(15.33)	(16.99)	(14.27)	(11.79)	(43.08)	(15.80)	(20.77)	(17.96)	(15.97)	(8.55)
Observations	2880	2880	2880	2880	2880	2880	2880	2880	2880	2880	2880	2880
R-Squared	0.000	0.000	0.000	0.000	0.000	0.035	0.000	0.004	0.003	0.004	0.005	0.017
Panel B: Dividend												
SURP	−0.014	−0.010	−0.015	−0.009	−0.013	−0.045	0.039	0.005	0.002	0.016	−0.003	−0.134
	(−0.99)	(−0.73)	(−1.05)	(−0.68)	(−0.94)	(−1.57)	(0.20)	(0.03)	(0.01)	(0.08)	(−0.02)	(−0.53)
H-TONE$_2$		−0.773						7.113				
		(−0.90)						(0.93)				
INTRO H-TONE$_2$			0.150		0.722	0.858			5.089		3.411	5.932
			(0.26)		(1.15)	(1.42)			(0.88)		(0.56)	(0.93)
Q&A H-TONE$_2$				−1.445*	−1.827**	−1.865**				7.159	5.356	7.101
				(−1.68)	(−1.98)	(−2.10)				(1.13)	(0.82)	(1.05)
INTRO COUNT						0.046						−0.733
						(0.58)						(−1.08)
Q&A COUNT						0.276***						−0.361
						(3.10)						(−0.66)
SIZE						−0.348***						−0.535

(*continued on next page*)

Table 8 (*continued*)

	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (−1,1)	CAV (2,60)	CAV (2,60)	CAV (2,60)	CAV (2,60)	CAV (2,60)	CAV (2,60)
						(−2.60)						(−0.90)
BM						−1.228***						4.379
						(−2.64)						(1.12)
ROA						0.041						0.384
						(1.10)						(1.33)
LEVERAGE						−0.016**						0.152***
						(−2.15)						(6.08)
VOLATILITY						0.413***						4.954**
						(2.75)						(2.40)
ANALYST						−0.101						−1.065
						(−0.90)						(−1.01)
DECLARATION						0.434						4.089
						(1.34)						(1.18)
Constant	5.939***	6.422***	5.844***	6.828***	6.604***	8.395***	71.125***	66.681***	67.892***	66.719***	65.662***	52.537***
	(33.66)	(11.82)	(15.04)	(12.55)	(11.59)	(5.94)	(44.63)	(13.97)	(18.48)	(15.97)	(13.89)	(6.05)
Observations	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440
R-Squared	0.000	0.001	0.000	0.003	0.004	0.065	0.000	0.001	0.001	0.001	0.001	0.027
Panel C: NonDiv												
SURP	0.029	0.028	0.028	0.030	0.030	0.014	0.161	0.141	0.140	0.199	0.179	0.093
	(0.63)	(0.61)	(0.60)	(0.66)	(0.63)	(0.25)	(0.74)	(0.62)	(0.62)	(0.91)	(0.81)	(0.37)
H-TONE$_2$		1.047						32.413***				
		(0.75)						(3.50)				
INTRO H-TONE$_2$			0.581		0.259	0.619			19.961***		11.957	12.932*
			(0.55)		(0.21)	(0.52)			(2.89)		(1.61)	(1.77)
Q&A H-TONE$_2$				1.045	0.929	1.015				28.482***	23.107***	22.605***
				(1.01)	(0.77)	(0.88)				(3.91)	(2.87)	(2.90)
INTRO COUNT						0.125						−0.389
						(0.97)						(−0.50)
Q&A COUNT						0.269***						−0.422
						(3.28)						(−1.00)
SIZE						−0.533***						−0.542
						(−3.48)						(−0.53)
BM						−0.030						−0.125
						(−1.11)						(−0.48)
ROA						0.081*						0.345
						(1.94)						(1.08)
LEVERAGE						−0.001						0.055
						(−0.12)						(1.24)
VOLATILITY						−0.107						−0.433
						(−0.72)						(−0.37)
ANALYST						0.067						−1.996
						(0.64)						(−1.38)
DECLARATION						−0.027						21.970***
						(−0.05)						(2.88)
Constant	6.325***	5.650***	5.946***	5.660***	5.565***	7.688***	68.971***	48.039***	55.928***	50.834***	46.444***	54.656***
	(52.62)	(6.58)	(8.76)	(9.61)	(7.93)	(8.43)	(37.72)	(7.99)	(12.18)	(10.36)	(8.70)	(5.07)
Observations	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440	1440
R-Squared	0.001	0.002	0.001	0.002	0.002	0.029	0.000	0.009	0.006	0.009	0.011	0.022

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

Table A-1
Harvard dictionary categories, description, and word counts.

Select Harvard IV-4 Psychosocial Dictionary categories	Number of words and description
Positive	1915 words of positive outlook
Negative	2291 words of negative outlook
Strong	1902 words implying strength
Weak	755 words implying weakness
Active	2045 words implying an active orientation
Passive	911 words indicating a passive orientation
Overstated	696 words indicating emphasis in realms of speed, frequency, causality, inclusiveness, quantity or quasi-quantity, accuracy, validity, scope, size, clarity, exceptionality, intensity, likelihood, certainty and extremity
Understated	319 words indicating de-emphasis and caution in these realms

current analysis is generally limited to the incremental disclosures reflected in the tone of the question and answer session as a whole, using presently available dictionaries. However, given the importance of the question and answer portion of conference calls, parsing the dialogue by individual call participants (i.e. managers and analysts) could further our understanding of the implications of tone in manager-investor communications. It could also help us discern any material differences in tone between the individual managerial roles (e.g. CEO and CFO). Additionally, given the stronger results for the context-specific, custom word categorization using the Henry (2008) dictionary, future research could attempt to build additional, more nuanced custom word lists beyond the simple positive and negative categories. For example, the Harvard dictionary categories of active and passive, strong and weak, and overstated and understated suggest a few such dimensions into which domain relevant custom dictionaries could be built. Passing a more finely toothed comb, so to speak, through the conference call text could help further our understanding of which specific tonal components convey the most useful information.

The present study is also limited to the examination of textual tone as a means of disclosure. However, as data becomes more readily available in additional forms, namely conference call audio and video files, the question and answer portion of conference calls may be further analyzed. This could include investigating other communication channels through which meaning is conveyed, such as speech tones, gestures, facial expressions, and so forth.

Table A-2
Henry word list.

Positive tone			Negative tone		
Above	Expands	Rewarded	Below	Fell	Weakness
Accomplish	Expansion	Rewarding	Challenge	Hurdle	Weaknesses
Accomplished	Good	Rewards	Challenged	Hurdles	Worse
Accomplishes	Greater	Rise	Challenges	Least	Worsen
Accomplishing	Greatest	Risen	Challenging	Less	Worsening
Accomplishment	Grew	Rises	Decline	Low	Worsens
Accomplishments	Grow	Rising	Declined	Lower	Worst
Achieve	Growing	Rose	Declines	Lowest	
Achieved	Grown	Solid	Declining	Negative	
Achievement	Grows	Strength	Decrease	Negatives	
Achievements	Growth	Strengthen	Decreased	Obstacle	
Achieves	High	Strengthened	Decreases	Obstacles	
Achieving	Higher	Strengthening	Decreasing	Penalties	
Beat	Highest	Strengthens	Depressed	Penalty	
Beating	Improve	Strengths	Deteriorate	Risk	
Beats	Improved	Strong	Deteriorated	Risks	
Best	Improvement	Stronger	Deteriorates	Risky	
Better	Improvements	Strongest	Deteriorating	Shrink	
Certain	Improves	Succeed	Difficult	Shrinking	
Certainty	Improving	Succeeded	Difficulty	Shrinks	
Definite	Increase	Succeeding	Disappoint	Shrunk	
Deliver	Increased	Succeeds	Disappointed	Slump	
Delivered	Increases	Success	Disappointing	Slumped	
Delivering	Increasing	Successes	Disappointment	Slumping	
Delivers	Larger	Successful	Disappoints	Slumps	
Encouraged	Largest	Up	Down	Smaller	
Encouraging	Leader		Downturn	Smallest	
Enjoy	Leading		Drop	Threat	
Enjoyed	More		Dropped	Threats	
Enjoying	Most		Dropping	Uncertain	
Enjoys	Opportunities		Drops	Uncertainty	
Exceed	Opportunity		Fail	Under	
Exceeded	Pleased		Failing	Unfavorable	
Exceeding	Positive		Fails	Unsettled	
Exceeds	Positives		Failure	Weak	
Excellent	Progress		Fall	Weaken	
Expand	Progressing		Fallen	Weakened	
Expanded	Record		Falling	Weakening	
Expanding	Reward		Falls	Weakens	

Acknowledgements

We are grateful to Ike Mathur (editor) and an anonymous referee. We also express our thanks to Alexander Wagner, Xu Li, C.F. Sirmans, Dean Gatzlaff, and seminar participants at Utah State University and the FMA annual conference for helpful comments and suggestions. All remaining errors are our own.

Appendix A. Appendix

The categories of the Harvard IV-4 Psychosocial Dictionary, which is the default dictionary for the General Inquirer software package, are far more expansive than just the two that have been incorporated into recent finance studies (positive and negative words). The broad dictionary covers such diverse areas as political, religious, academic, gender specific, or even race related words. We identified eight relevant categories for the purpose of establishing an overall measure of positive and negative tone. Category descriptions, the number of root words, and a corresponding description of the type of words found in each category are in Table A-1.

One of the limitations of applying the general Harvard dictionary is the potential misclassification of words due to context specific meanings (Loughran and McDonald, 2011). For example, the Harvard dictionary categorizes *Volatility* as a negative word, which is correct in many settings. However, option traders view *Volatility* positively. In light of such limitations, content analysis should be viewed as a fairly blunt instrument. One way to overcome some misclassification is through the use of custom dictionaries.

Similar to Loughran and McDonald (2011) and Henry (2008) argues that the Harvard IV-4 Psychosocial Dictionary is too broad to be properly applied within the realm of finance and related disciplines. As such she creates her own list of "domain relevant" words that connote either a positive or negative outlook using earnings announcement specific terminology. The list is shown in Table A-2. Of the collection of 199 words in the Henry list, all are found within the Harvard dictionary categorized under one of the eight categories listed in Table A-1, except four words – *Deteriorate*, *Downturn*, *Hurdle*, and *Hurdles*. We programmed the Henry word list into the General Inquirer as a custom dictionary which still allowed us to take advantage of the GI's powerful disambiguation (word recognition) capacity.

The GI applies over 6300 disambiguation rules in order to properly identify words within a body of text. For example, the program is able to identify *Accomplished*, *Accomplishes*, and *Accomplishing* as having the same common root, Accomplish, which it then classifies according to the definitions or categorizations of the dictionary applied to the analysis. The GI will recognize *Accomplishments*, as a separate word with the root *Accomplishment*, which then receives its own categorization based on the dictionary.

References

Abarbanell, J.S., Bernard, V.L., 1992. Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. Journal of Finance 47, 1181–1207.
Akbas, F., Boehmer, E., Erturk, B., Sorescu, S., 2010. Don't shoot the messenger: short interest, returns, and fundamentals. Working paper, Texas A&M University.
Allen, F., Michaely, R., 2002. Payout policy. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), Handbook of the Economics of Finance. Elsevier Science, Amsterdam.
Amir, E., Lev, B., 1996. Value-relevance of non-financial information: the wireless communications industry. Journal of Accounting and Economics 22, 3–30.
Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159–178.
Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21 (2), 785–818.
Beaver, W.H., 1968. The information content of annual earnings announcements. Journal of Accounting Research 6, 67–92.
Berelson, B.R., 1952. Content Analysis in Communication Research. Free Press, New York.
Bernard, V., 1992. Stock price reactions to earnings announcements: a summary of recent anomalous evidence and possible explanations. In: Thaler, R. (Ed.), Advances in Behavioral Finance. Russell Sage Foundation, New York.
Bernard, V., Thomas, J., 1989. Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting Research 27, 1–36.
Bernard, V., Thomas, J., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305–340.
Brav, A., Graham, J., Harvey, C., Michaely, R., 2005. Payout policy in the 21st century. Journal of Financial Economics 77, 483–527.
Brav, A., Heaton, J.B., 2002. Competing theories of financial anomalies. Review of Financial Studies 15 (2), 575–606.
Brown, L.D., Hagerman, R.L., Griffin, P.A., Zmijewski, M.E., 1987. An evaluation of alternative proxies for the market's assessment of unexpected earnings. Journal of Accounting and Economics 9, 159–193.
Brown, S., Hillegeist, S.A., Lo, K., 2004. Conference calls and information asymmetry. Journal of Accounting and Economics 37, 343–366.
Campbell, J.Y., Ramadorai, T., Schwartz, A., 2009. Caught on tape: institutional trading, stock returns, and earnings announcements. Journal of Financial Economics 92, 66–91.
Chung, C.K., Pennebaker, J.W., 2007. The psychological functions of function words. In: Fiedler, K. (Ed.), Social Communication. Psychology Press, New York.
Collins, D.W., Li, O.Z., Xie, H., 2009. What drives the increased informativeness of earnings announcements over time? Review of Accounting Studies 14, 1–30.
Core, J.E., 2001. A review of the empirical disclosure literature: discussion. Journal of Accounting and Economics 31, 441–456.
Davis, A.K., Piger, J.M., Sedor, L.M., 2008. Beyond the numbers: managers' use of optimistic and pessimistic tone in earnings press releases. AAA 2008 Financial Accounting and Reporting Section (FARS) Paper.
DeAngelo, H., DeAngelo, L., 2006. The irrelevance of the MM dividend irrelevance theorem. Journal of Financial Economics 79, 293–315.
DeAngelo, H., DeAngelo, L., Skinner, D.J., 2004. Are dividends disappearing? Dividend concentration and the consolidation of earnings. Journal of Financial Economics 72, 425–456.
Demers, E., Vega, C., 2008. Soft information in earnings announcements: news or noise? Working paper, Federal Reserve Board.
Denis, D.J., Osobov, I., 2008. Why do firms pay dividends? International evidence on the determinants of dividend policy. Journal of Financial Economics 89 (1), 62–82.
Diefenbach, D.L., 2001. Historical foundations of computer-assisted content analysis. In: West, M.D. (Ed.), Theory, Method, and Practice in Computer Content Analysis. Ablex, New York.
Doyle, J.T., Lundholm, R.J., Soliman, M.T., 2003. The predicative value of expenses excluded from pro forma earnings. Review of Accounting Studies 8, 145–174.
Dye, R.A., 2001. An evaluation of "essays on disclosure" and the disclosure literature in accounting. Journal of Accounting and Economics 32, 181–235.
Engelberg, J., 2008. Costly information processing: evidence from earnings announcements. Working paper, Northwestern University.
Foster, G., Olsen, C., Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. The Accounting Review 59, 574–603.
Francis, J., Lafond, R., Olsson, P., Schipper, K., 2007. Information uncertainty and post-earnings-announcement-drift. Journal of Business Finance & Accounting 34, 403–433.
Frankel, R.M., Johnson, M., Skinner, D.J., 1999. An empirical examination of conference calls as a voluntary disclosure medium. Journal of Accounting Research 37, 133–150.
Frankel, R.M., Mayew, W.J., Sun, Y., 2010. Do pennies matter? Investor relations consequences of small negative earnings surprises. Review of Accounting Studies 15, 220–242.
Graham, J.R., Harvey, C.R., Rajgopal, S., 2005. The economic implications of corporate financial reporting. Journal of Accounting and Economics 40, 3–73.
Hart, R.P., 2001. Redeveloping DICTION: theoretical considerations. In: West, M.D. (Ed.), Theory, Method, and Practice in Computer Content Analysis. Ablex, New York.
Healy, P.M., Palepu, K.G., 2001. Information asymmetry, corporate disclosure, and the capital markets: a review of the empirical disclosure literature. Journal of Accounting and Economics 31, 405–440.
Henry, E., 2008. Are investors influenced by how earnings press releases are written? The Journal of Business Communication 45 (4), 363–407.
Hoberg, G., Prabhala, N.R., 2009. Disappearing dividends, catering, and risk. Review of Financial Studies 22 (1), 79–116.
Holsti, O.R., 1963. Computer content analysis. In: North, R.C., Holsti, O.R., Zaninovich, M.G., Zinnes, D.A. (Eds.), Content Analysis: A Handbook with Applications for the Study of International Crisis. Northwestern University, Evanston, IL.
Irani, A.J., 2004. The effect of regulation fair disclosure on the relevance of conference calls to financial analysts. Review of Quantitative Finance and Accounting 22, 15–28.
Kimbrough, M.D., 2005. The effect of conference calls on analyst and market underreaction to earnings announcements. The Accounting Review 80, 189–219.
Krippendorf, K., 2004. Content Analysis: An Introduction to its Methodology. Sage Publications Inc., Thousand Oaks, CA.
Laffal, J., 1970. Toward a conceptual grammar and lexicon. Computers and the Humanities 4 (3), 173–186.

Lang, M.H., Lundholm, R.J., 2000. Voluntary disclosure and equity offerings: reducing information asymmetry or hyping the stock? Contemporary Accounting Research 17 (4), 623–662.
Lewellen, J., Shanken, J., 2002. Learning, asset-pricing tests, and market efficiency. Journal of Finance 57, 1113–1145.
Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics 45, 221–247.
Liang, L., 2003. Post-earnings announcement drift and market participants' information processing biases. Review of Accounting Studies 8, 321–345.
Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. Journal of Finance 66 (1), 35–65.
Matsumoto, D., Pronk, M., Roelofsen, E., 2007. Conference call disclosures and firm performance: an empirical analysis of the length and content of earnings-related conference calls. Working paper, University of Washington.
Mayew, W.J., Venkatachalam, M., forthcoming. The power of voice: Managerial affective states and future firm performance. Journal of Finance.
Mehrabian, A., 1971. Silent Messages. Wadsworth Publishing, California.
Mehrabian, A., 1972. Nonverbal Communication. Aldine Atherton, Illinois.
Merkl-Davies, D.M., Brennan, N.M., 2007. Discretionary disclosure strategies in corporate narratives: incremental information or impression management? Journal of Accounting Literature 26, 116–194.
Miller, M., Modigliani, F., 1961. Dividend policy, growth, and the valuation of shares. Journal of Business 34, 411–433.
NIRI, 1996. Standards and guidance for disclosure. National Investor Relations Institute, Vienna, VA.
NIRI, 2004. Standards of practice for investor relations. National Investor Relations Institute, Vienna, VA.
Pennebaker, J.W., Mehl, M.R., Niederhoffer, K.G., 2003. Psychological aspects of natural language use: OUR words, our selves. Annual Review of Psychology 54, 547–577.
Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Review of Financial Studies 22 (1), 435–480.
Petty, R.E., Cacioppo, J.T., 1986. Communication and Persuasion: Central and Peripheral Routes to Attitude Change. Springer-Verlag, New York.
Psathas, G., 1969. The general inquirer: useful or not? Computers and the Humanities 3 (3), 163–174.
Sadique, S., In, F., Veeraraghavan, M., 2008. The impact of spin and tone on stock returns and volatility: evidence from firm-issued earnings announcements and the related press coverage. Working paper, Monash University.
Stone, P.J., Dunphy, D.C., Smith, M.S., Ogilvie, D.M., 1966. The General Inquirer: A Computer Approach to Content Analysis. The MIT Press, Cambridge, MA.
Sunder, S.V., 2002. Investor access to conference call disclosures: Impact of regulation fair disclosure on information asymmetry. Working paper, New York University.
Tetlock, P.C., 2007. Giving content to investor sentiment: the role of media in the stock market. Journal of Finance 62, 1139–1168.
Tetlock, P.C., Saar-Tsechansky, M., Macskassy, S., 2008. More than words: quantifying language to measure firms' fundamentals. Journal of Finance 63, 1437–1467.
Uang, J.Y., Citron, D.B., Sudarsanam, S., Taffler, R.J., 2006. Management going-concern disclosures: impact of corporate governance and auditor reputation. European Financial Management 12, 789–816.
Verrecchia, R.E., 2001. Essays on disclosure. Journal of Accounting and Economics 32, 97–180.
West, M.D., Fuller, L.K., 2001. Toward a typology and theoretical grounding for computer content analysis. In: West, M.D. (Ed.), Theory, Method, and Practice in Computer Content Analysis. Ablex, New York.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity. Econometrica 48, 817–838.
Zhang, X.F., 2006. Information uncertainty and stock returns. Journal of Finance 61, 105–136.