The Power of Voice: Managerial Affective States and Future Firm Performance

WILLIAM J. MAYEW and MOHAN VENKATACHALAM*

ABSTRACT

We measure managerial affective states during earnings conference calls by analyzing conference call audio files using vocal emotion analysis software. We hypothesize and find that, when managers are scrutinized by analysts during conference calls, positive and negative affects displayed by managers are informative about the firm's financial future. Analysts do not incorporate this information when forecasting near-term earnings. When making stock recommendation changes, however, analysts incorporate positive but not negative affect. This study presents new evidence that managerial vocal cues contain useful information about a firm's fundamentals, incremental to both quantitative earnings information and qualitative "soft" information conveyed by linguistic content.

> It is not what you say that matters but the manner in which you say it; there lies the secret of the ages.
>
> —William Carlos Williams

MANAGERS DISSEMINATE AN ABUNDANT amount of quantitative and qualitative information about their actions and firm performance on both a voluntary and a mandatory basis through several avenues, including press releases, quarterly and annual reports, shareholder meetings, and earnings conference calls. Prior literature is replete with studies that evaluate the extent to which capital market participants react to quantitative information contained in these disclosures. Only recently have researchers begun to explore the capital market implications of *qualitative* verbal communication via financial news

stories (Tetlock (2007), Tetlock, Saar-Tsechansky, and Macskassy (2008)), annual reports (Feldman et al. (2010), Loughran and McDonald (2011)), conference presentations (Bushee, Jung, and Miller (2011)), and earnings press releases (Davis, Piger, and Sedor (2011), Demers and Vega (2010)). In general, the findings support the hypothesis that qualitative verbal communication by managers is incrementally useful to quantitative information in predicting future firm fundamentals and stock returns. This paper extends this line of inquiry by focusing on how one important type of *nonverbal* communication, vocal cues from executives during conference calls, can inform investors about a firm's future profitability and stock returns.

Using a sample of conference call audio files and commercially available Layered Voice Analysis (LVA) software, we analyze managerial vocal cues to measure positive and negative dimensions of a manager's affective or emotional state. Research in linguistics and social psychology has long recognized that the human voice conveys considerable information over and above the literal meaning contained in verbal content (Caffi and Janney (1994)). Vocal cues or expressions are considered important in drawing inferences about both positive affective states (e.g., happiness, excitement, and enjoyment) and negative affective states (e.g., fear, tension, and anxiety). The appraisal theory of emotion suggests that affective states arise from an individual's cognitive evaluation of a situation or stimulus and its attendant implications for personal well-being. In other words, affective states are responses to interpretation and evaluation of events and stimuli and hence reveal useful information. The extent of the emotional response will be a function of the strength of the stimulus or elicitor.

In the context of conference calls, the external stimulus that is likely to produce affective states is the questioning by analysts during the conference call. Further, the affective state is likely to be more prominent when the analysts' questions are more pointed and scrutinizing. Consequently, affective states elicited from analysts' probing during the conference call are likely to contain useful information about the firm's economic activities and performance. Survey evidence by Graham, Harvey, and Rajgopal (2005) suggests that managers who miss analysts' earnings expectations face extensive questioning during the conference call. We therefore posit that affective states are most likely to be elicited during the question and answer portion of the conference call, and, in particular, when firms have missed earnings expectations and are subject to intense scrutiny by analysts.

If affective states exhibited by managers during conference calls contain new information about firm fundamentals, we expect investors to incorporate this information into stock prices. Consistent with this prediction, we find—even after controlling for the linguistic content in the conference calls—that both positive and negative affects exhibited by managers during the question and answer portion of earnings conference calls are associated with contemporaneous stock returns. Moreover, the stock market's response to the information contained in the affective state is more pronounced when managers are "interrogated" and subject to more scrutiny during the conference calls.

While investors react to affective states as if they carry value relevant information, analysts do not react in a similar fashion when forecasting near-term earnings. That is, we are unable to document a relation between affective states and forecast revision magnitudes of one-quarter-ahead earnings following the conference call. This result is open to two interpretations. Either analysts fail to appreciate the valuation implications of nonverbal cues or analysts do consider this information but incorporate it as part of the "soft" information in determining long-term forecasts that underpin stock recommendations. Our evidence is consistent with the latter interpretation. We find a positive association between positive affect and changes in stock recommendations immediately following the call. However, we find no association between negative affect and recommendation changes, a finding that is perhaps consistent with analyst incentives to delay incorporating bad news into their stock recommendations (McNichols and O'Brien (1997), O'Brien, McNichols, and Lin (2005)).

Next, we examine whether the stock market reaction around the earnings call is consistent with future firm-specific information about fundamentals. We find that both positive and negative affects are associated with future unexpected earnings (based on analyst expectations) measured over the two subsequent quarters. We also examine firm-issued press releases from news wires over the 180 days following the conference call. We classify news releases as good or bad depending on the market reaction surrounding the press release and compute the proportion of bad news releases following the conference call. Our findings suggest that managers who exhibit positive affect issue a lower proportion of bad news press releases in the future.

Finally, we examine whether market participants reflect managerial affect for future performance with any delay. We find that negative affect is related to cumulative abnormal returns over the subsequent 180 trading days following the earnings conference call. We cannot identify for certain why market participants fail to incorporate negative affect completely. One plausible explanation is that market participants follow analysts' recommendations, which do not completely take into account the information in negative affect. Additional analyses reveal that, when analysts observe negative affect, they are less likely to revise their outstanding earnings forecasts. Together, our evidence is not consistent with analysts' failure to incorporate negative affect; rather, it is more consistent with analysts' reluctance to revise forecasts and recommendations when faced with "soft" negative information about a firm's future prospects (McNichols and O'Brien (1997), O'Brien et al. (2005)). Regardless, we caution the reader that this apparent underreaction does not imply a plausible trading strategy as transaction costs could eliminate any potential trading profits.

This study makes the following contributions. First, to our knowledge, this is the first paper to provide evidence on the role of nonverbal communication in a capital market setting. We apply findings in social psychology research that provide unequivocal support for vocal expressions as one particular type of nonverbal communication that is influential when communicating messages over and above their verbal content (Mehrabian and Weiner (1967), Scherer, London, and Wolf (1973), Scherer (2003)). Our findings confirm that important

information can be gleaned from vocal cues in the capital market setting by showing that managers' emotional state is associated with stock returns and future firm performance, after we control for quantitative information and qualitative verbal content. Future research in both economics and psychology can explore vocal cues in other settings. For example, examining information about the affective states of economic leaders like the Federal Reserve Chairman can perhaps be informative about broader changes in economic fundamentals.

Second, our results provide new insights into how conference calls can provide information to financial markets. Prior research documents that conference calls provide significant information to market participants above and beyond that contained in the earnings press release (Frankel, Johnson, and Skinner (1999)). As conference call audio broadcasts are commonplace for many firms and open to public access subsequent to Regulation FD, our findings suggest that investors can and do use vocal cues during such communication to learn about a manager's affective state and in turn about the firm's financial future.

Our study is subject to the following caveat. Although our evidence is consistent with the LVA software generating useful proxies for managerial affect in the capital market setting, the generalizability of our results largely depends on the validity of the LVA-based measures. We offer some preliminary evidence on the construct validity of the LVA measures that we use in this paper, but certainly more empirical validation of this software's reliability is warranted. We view our empirical results as complementary to recent experimental investigations of the construct validity of LVA metrics in various settings (Elkins (2010), Elkins and Burgoon (2010), Han and Nunes (2010), Hobson, Mayew, and Venkatachalam (2011)).

The paper proceeds as follows. In Section I, we review related literature and develop our hypotheses. Section II discusses the nonverbal measures used in the study. In Section III, we outline our sample selection, define our variables of interest, and provide descriptive statistics. Sections IV and V discuss our empirical results and additional analyses, and in Section VI we offer concluding remarks.

I. Related Research and Hypothesis Development

A. Related Research

Social psychology research suggests that nonverbal cues such as vocal and facial expressions influence how a message is interpreted. Communication experts generally agree that in face-to-face conversations, only a small fraction of the message regarding emotional state is contained in the verbal content (Mehrabian (1971)). A significant component of the message is contained in vocal attributes such as voice intonation, accent, speed, volume, and inflection. Kinesics—that is, facial expressions, postures, and gestures—also plays a large role in communication. However, we do not study these traits in this paper and therefore do not elaborate further on the role of kinesics. We instead focus

on the vocal channel and describe how voice can convey emotions or affective states reliably to a receiver (Juslin and Laukka (2003)).[1]

The expression and perception of emotional states via vocal cues are fundamental aspects of human communication. People express emotions by yelling; using a quiet, low, or monotonous voice; and, at the extreme, by being silent (Walbott, Ricci-Bitti, and Banninger-Huber (1986)). Such expression of emotions through voice can be used to convey information or influence others. Several studies have shown that the tone of a person's voice signals information about an affective state that is not revealed by the verbal content or facial expressions associated with the message (Zuckerman et al. (1982)). Juslin and Scherer (2005) review 50 years of research establishing that acoustic voice patterns provide insights into the speaker's affective, or emotional, state. While the role of nonverbal cues has been studied extensively in the social psychology literature, it is virtually absent from the accounting and finance literatures.

Corporate financial reporting represents an important channel for managers to communicate information to various stakeholders, and much of the literature focuses primarily on the capital market implications of quantitative information disclosed in the financial statements. Recently, researchers have begun to explore verbal communication as an additional mechanism through which information is conveyed and used in capital markets. For example, the informativeness of verbal communication has been documented in the context of financial news stories (Tetlock (2007), Tetlock et al. (2008)) and messages in Internet chat rooms (Antweiler and Frank (2004)). Extending these findings to written firm communications, research finds evidence of value-relevant information in the linguistic narratives of earning press releases (Davis et al. (2011), Demers and Vega (2010)) and mandatory regulatory filings (Feldman et al. (2010), Loughran and McDonald (2011)). Voluntary communications during presentations by corporate executives at investor conferences have also been shown to convey important information (Bushee et al. (2011)).

While this growing body of literature explores the role of verbal communication in the financial markets arena, the implications of nonverbal communication represent a fairly nascent and uncharted territory. One exception is Coval and Shumway (2001), who examine the role of ambient noise level in the Chicago Board of Trade's bond futures trading pit. They find that ambient sound level conveys economically and statistically meaningful information and that traders process subtle and complex nontransaction signals in determining equilibrium prices. While this finding suggests that decibel levels in trading pits have information content for equilibrium supply-and-demand conditions in the futures market, it does not speak directly to the specific attributes of nonverbal communication between managers and market participants that we address in our study.

---

[1] Although we use the terms affect and emotion interchangeably, there is a subtle but important difference between the two. Emotion refers to a feeling that occurs in response to events, while affect is viewed as a valence of an emotional state (Frijda (1993)).

B. Hypotheses

An individual's affective or emotional state allows us to draw inferences about the events or type of events that caused an individual to be in such a state. These inferences are based on the appraisal theory of emotion, which is founded on the notion that emotions arise or are elicited by evaluations or appraisals of events and situations (Arnold (1960), Roseman (1984), Lazarus (1991)). For example, a positive state is elicited by a successful outcome such as winning a basketball game, passing an exam, or being admitted to a prestigious university. In contrast, a negative state is elicited by personal loss, frustration, cognitive dissonance, or simply a bad outcome. Frijda (1988) uses the term the Laws of Situational Meaning and Concern and states that "emotions arise in response to the meaning structures of given situations; . . . . arise in response to events that are important to the individual's goals, motives, and concerns" (pp. 349, 351). In other words, emotions arise from an individual's cognitive evaluation and interpretation of events and situations that in turn have implications for personal well-being.

While human emotions can arise without an external stimulus, most emotions are the result of social and interpersonal communication (Andersen and Guerrero (1998)). The triggering event can be external, such as a loud noise, or internal, such as a physiological change. External elicitors invoke cognitive processes that in turn trigger certain affective states. Most extant research in psychology focuses on external stimuli because of the difficulties in identifying internal elicitors that trigger affective states (Lewis (1993)). In order for an emotional state to arise, some event acts as a stimulus that in turn triggers a change in the state of the individual.

In the context of financial markets, in which managers communicate information to investors about both past and future performance, it is likely that managers exhibit different affective states depending on their interpretation of events and situations pertaining to the firm. Such affective states are most likely elicited when managers answer analyst questions. The determination of managerial affective states should allow investors to infer the managers' implicit assessment of firm performance, both past and future. For example, a manager is likely to exhibit positive affect during analyst questioning if the manager expects positive future firm performance due to private information regarding current outcomes (e.g., persistence of current-period earnings) and/or future outcomes (e.g., prospective drug approval, anticipated orders, successful outcome of strategic initiatives such as restructuring). In such instances, a manager is more likely to be excited or exhibit positive psychological arousal in communication with investors.[2]

---

[2] We assume that a manager's affective state is not an innate characteristic of the manager per se. Rather, it is time-dependent and is a function of private information about the firm that managers possess during the conference call. It is plausible that a manager could exhibit both positive and negative affective states during the conference call if the manager has both good news and bad news about specific issues discussed during the conference call. For example, a manager may discuss poor past performance in the form of a negative earnings surprise and at the same time discuss better expected future performance as a result of an increasing backlog of orders.

In contrast, a manager may exhibit negative affect when possessing negative private information. Examples include information about the transitory nature of accounting earnings, impending lawsuits, product failures, or order cancellations. Negative affect may also stem from managers' psychological discomfort due to cognitive dissonance. The theory of cognitive dissonance, developed by Festinger (1957), is based on the notion that inconsistency between an individual's beliefs and actions creates a feeling of discomfort and anxiety. In experiments conducted by Elliot and Devine (1994), counterattitudinal behavior evoked psychological discomfort, arousing a negatively valenced state (see also Harmon-Jones (2000)).

To apply cognitive dissonance in the economic setting we explore here, consider a manager who believes that she is competent and in control of the firm she operates. Information about firm performance would reflect her actions taken while running the firm. If the manager has private information that is inconsistent with her own beliefs regarding her competence, an uncomfortable emotional state will arise from this dissonance. As such, we posit that cognitive dissonance–induced negative affect should be indicative of potential bad news or uncertainty about good news. Therefore, if we observe a manager in a negative affective state, it is more likely that events and circumstances are unfavorable and/or that the manager is psychologically uncomfortable due to cognitive conflicts in elements of information that the manager has.

If positive (negative) managerial affect is reflective of favorable (unfavorable) private information, we should observe a positive (negative) capital market response surrounding the communication date. Observing such a market response is contingent on (1) the strength of the stimulus that generates the affective state, which in this setting is the intensity of analyst probing during the conference call, and (2) the efficiency with which market participants observe and act on the information contained in the affective state.

Extant psychology and emotion research suggests that both conditions are likely to be satisfied. Research in social psychology suggests that vocal indicators of various emotions are accurately detected and are often as good as or better than those of facial cues and expressions (Kappas, Hess, and Scherer (1991)). It is also widely accepted that one's voice is not easily controlled and that the voice channel "leaks" more information than facial cues (Ekman and Friesen (1974)). Evidence in Ambady and Rosenthal (1993) suggests that human beings can form impressions and judgments from even "thin slices" of nonverbal behavior. Emotional contagion research (Hatfield, Cacioppo, and Rapson (1994), Neumann and Strack (2000)) suggests that the perception by the receiver of another person's behavior might activate the same cognitive processes in the receiver that generated the other person's behavior. In other words, affective states are transferred between individuals either consciously by imitation or subconsciously. Barsade (2002) extends this research to show that emotional contagion occurs not only from one individual to another but also from one individual to a group. Hence, the CEO's emotional expression during conference calls may evoke congruent feelings in the analysts and investors who listen to the CEO. Nevertheless, documenting a statistical association between affective

states and capital market responses relies heavily on the precision with which affect is empirically measured.

II. Measuring Nonverbal Communication

The main challenge in this study is to construct useful and reliable measures of affective states from nonverbal communication by firm managers. We use CEO and CFO voice recordings from earnings conference calls to develop measures of managers' emotive states when communicating information to analysts and investors.[3] There are several advantages to using the audio content in earnings conference calls. First, conference calls represent a common and important disclosure mechanism for U.S. firms as most public firms regularly host quarterly earnings conference calls (Skinner (2003)). A second advantage of using conference calls is that, unlike annual meetings where managers appear face-to-face to meet with current investors, conference calls offer one of the few opportunities for firms to communicate directly with current and potential investors as well as other stakeholders. Third, because conference calls are rarely broadcast over video, other channels of nonverbal communication such as facial expressions and gestures do not contaminate the signal in the voice channel. In other words, we are able to isolate the vocal channel of the nonverbal communication.[4] Lastly, as a practical matter, we are able to obtain audio files of the conference calls from the Thomas Reuters StreetEvents database.

We construct measures of affective states with the help of a computer software program that uses LVA technology. LVA was invented in 1997 by Nemesysco Ltd. in Israel. LVA is comprised of a set of proprietary signal processing algorithms that extract and combine attributes from the voice in order to identify different types of stress, cognitive processes, and emotional reactions. The software performs analysis and provides output at the voice segment level. A voice segment is a logical portion of continuous voice (one word to a few words) that may range in length from $4/10^{th}$ of a second to 2 seconds. The original objective of the LVA technology was to measure several different emotions that, in combination, would enable a user to conclude whether a speech segment was at low risk or high risk of being deceptive. To that end, LVA-based software creates output based on different layers of analysis. The base layer extracts and combines raw vocal attributes, the next layer creates fundamental emotion variables, and the final layer creates conclusion variables that result from combining results from prior layers. The LVA technology underpins various software products for commercial purposes (see

---

[3] In the psychology literature, nonverbal cues are often generated by using actors to produce vocal emotion expressions, and human judges are used as "decoders" to determine whether such vocal patterns are recognized. While professional actors can provide strong vocal cues and it is easy to get consistent audio recordings, their emotional portrayals may not be ecologically valid and therefore differ from vocal expressions that occur in real life.

[4] The message recipient may react to the verbal or linguistic aspect of the communication in addition to or in lieu of the nonverbal content. We control for this possibility by including linguistic tone in the empirical analysis.

www.nemesysco.com/solutions.html for a complete list) and, depending on the particular software application, the specific output from each layer varies.[5] We use the LVA-based Ex-Sense Pro-R (version 4.3.9) Digital Emotion Analyzer application because it is purportedly designed for business applications and because it is most cost effective given our research objective.[6]

In the LVA software we use, the base layer variables (technically termed SPT, SPJ, JQ, and AVJ) are raw values obtained from unique measurements of the vocal wave. In addition to the raw values, a set of parallel calibration values (calSPT, calSPJ, calJQ, and calAVJ) are derived from "emotion free" voice segments. These segments occur during the beginning of a conversation, at which time the general baseline emotional state of the tested subject is presumed to be present. Differencing off the subject-specific calibrated value of each raw base layer variable provides the four fundamental variables of the LVA software we use: Emotion Level, Cognition Level, Global Stress, and Thinking Level. Adjusting for baseline values is of critical importance in the LVA analysis so that the system can take into account different emotional states and different personality structures, as well as acoustic and audio quality issues. Emotion Level purports to capture excitement. Cognition Level purports to capture cognitive dissonance. Global Stress purports to capture physical arousal and alertness, and Thinking Level purports to capture the mental effort behind what the subject is saying.

In addition to these four fundamental variables, the software also provides "conclusion" variables (also known as algorithmic values), which are proprietary combinations of the four fundamental variables and base layer raw value variables. These conclusion variables (e.g., Lie Stress) are meant to allow a user to draw conclusions about whether a given speech segment should be further examined or treated as potentially untruthful. Since the objective of our study is not to detect lies, we do not use the conclusion variables produced by the software.

For our empirical analysis, we select the two measures implicit in the LVA fundamental variables that are relevant for operationalizing positive and negative managerial affects.[7] The first measure, Cognition Level, purportedly measures the level of cognitive dissonance (Festinger (1957)). Cognitive dissonance is the uncomfortable, anxious feeling an individual experiences when beliefs and actions are contradictory, leading to a negative affective state (Forgas

---

[5] For example, LVA 6.50 is the security level version of the software used for police interrogations and military operations. Ex-Sense Pro-R is a digital emotion analyzer marketed for business solutions such as interviewing customers, employees, and potential business partners. QA5 is designed for emotion detection in call center conversations. StressIndicator is marketed to individuals for managing stress in daily life as a health care application.

[6] Hereafter, when we refer to LVA we are referring to Ex-Sense Pro-R.

[7] We do not consider the two other fundamental variables, Global Stress and Thinking Level, because we are unable to posit directional predictions for a market response for these variables. For example, it is unclear whether an individual is physically aroused or alert for good news reasons or bad news reasons. Similarly, it is unclear whether extensive thinking, or a high cognitive load, means good news or bad news. Reestimating all empirical specifications after including these variables and the conclusion variables does not alter our inferences.

(2001)). Cognition Level from the LVA software takes on values ranging from 30 to 300, with values above 120 indicating abnormally high levels. Thus, assuming that Cognition Level captures cognitive dissonance, higher values indicate more cognitive dissonance, and in turn more negative affect (hereafter, *NAFF*).

The second measure, Emotion Level, purportedly measures the level of excitement exhibited by the subject. Excitement is one of the biological expressions that accompanies a positive affective state (Tomkins (1962)). As with Cognition Level, Emotion Level values range from 30 to 300. Emotional levels greater than 110 indicate abnormally high levels of excitement. Thus, the higher the emotion level, the larger the positive affect (hereafter, *PAFF*).[8]

Our decision to use LVA-based software results from a careful cost–benefit analysis on many dimensions. The trade-offs we consider are the parameters provided by the software, the monetary cost of the software, and, most important, the construct validity of the parameters. We employ LVA-based technology instead of other commercial voice stress analyzers (such as Psychological Stress Evaluator (PSE) or Computerized Voice Stress Analysis (CVSA)) for two reasons. First, they offer only basic speech segment diagnostics of true or false outcomes without providing variables that capture positive and negative emotions. Second, some of these softwares require enormous capital investments.

We choose a commercial product in LVA instead of constructing emotion metrics from vocal acoustic features directly because it is not clear from the literature which vocal emotion measurement model would be most appropriate. The literature on identifying which acoustic features to extract from voice and how to combine them for affective state classification is vast and evolving, with little agreement on which models are superior (Ververidis and Kotropoulos (2006), Wu, Yeh, and Chuang (2009), Schuller (2010), Yang and Lugger (2010)). Naturally, using a commercial product like LVA is also limiting because the developers are reluctant to divulge the specific acoustic features they extract from voice and how they combine these features. As a result, in Section V.A, we examine the association between the LVA metrics we use and common acoustic voice features used in the measurement of emotion (Owren and Bachorowski (2007)) to begin to bridge the gap between commercial products and the academic literature on emotion detection.

Regarding the construct validity of LVA metrics, we summarize the literature that examines LVA's performance.[9] Because the software was originally designed to detect deception, studies commonly obtain voice samples from truth tellers and liars in experiments or field studies and examine whether LVA algorithmic "conclusion" variables can successfully distinguish between truthful

---

[8] In the limit, a high emotion level is likely when the context is either deceptive or traumatic in nature. To the extent that such situations dominate in the determination of the *PAFF*, it will bias against finding the predicted relation.

[9] A more detailed summary of these individual studies and an overview of the literature on extracting emotion from voice are available in the Internet Appendix. (An Internet Appendix for this article is available online in the "Supplements and Datasets" section at http://www.afajof.org/supplements.asp.)

and deceptive speech segments.[10] Several studies document that LVA algorithmic metrics for detecting deception perform no better than chance levels.[11] Lacerda (2009) and Erikkson and Lacerda (2007) question the validity of LVA overall and suggest the lack of results in the literature pertaining to lie detection arise because (1) LVA does not extract relevant information from the speech signal and (2) variation in LVA output measures is simply an artifact of the digitization of analog speech signals.

Other research suggests it would be premature to dismiss LVA as invalid. More recent research relaxing reliance on the built-in algorithmic conclusion variables for identifying deception find that the LVA variables from more primitive layers do statistically discriminate between truth and deception (Elkins (2010), Elkins and Burgoon (2010)). These findings are similar to Brown et al. (2003), who perform exploratory logistic regression analysis for predicting deception and find that detection capabilities are greatly improved using more primitive LVA variables instead of the prepackaged algorithmic variables. Moreover, Elkins and Burgoon (2010) show that these more primitive LVA measures can distinguish between responses to charged and neutral questions, and that the full collection of primitive LVA measures appears to identify latent constructs that correlate with self-reported subject scores of emotional state. On the basis of this evidence, they conclude that LVA can discriminate vocal responses characterized by stressful and emotional tone.

Other research explores specific base layer values and fundamental LVA variables in isolation. Harnsberger et al. (2009) investigate whether the JQ base layer metric, which represents the uncalibrated Global Stress metric, is higher in settings in which electric shocks were administered during speaking versus settings in which no such shock was administered. They find little evidence that the JQ parameter can detect the stress associated with electric shocks at better than chance levels. In contrast, Konopka, Duffecy, and Hur (2010) find that the LVA Global Stress metric can discriminate among speech samples from Vietnam veterans diagnosed with posttraumatic stress disorder and those without such a diagnosis. Hobson et al. (2011) conduct an experiment that invokes cognitive dissonance from misreporting and document a strong association between subject cognitive dissonance levels and the LVA fundamental variable, Cognition Level. In the marketing literature, Han and

---

[10] Early research discussed in Palmatier (2005) compares LVA deception detection capabilities with those of the polygraph, and finds that LVA works better than chance and similarly to the polygraph. However, direct assessment of the LVA parameters is not possible due to the research design. In the study, real-life speech samples from police interrogations, where truth and deception were known with certainty, were independently sent to a polygraph examiner and an examiner trained in LVA. Conclusions about truth and deception were then submitted by both the polygraph and LVA examiner and compared with ground truth, making it impossible to isolate the predictive ability of the LVA metrics separately from the ability of the LVA examiner. In a similar research design, more recent research by Adler (2009) using sex offender speech samples also finds LVA to predict deception at rates similar to the polygraph.

[11] See, for example, Harnsberger et al. (2009), Damphousse et al. (2007), Sommers et al. (2007), Sommers (2006), Gamer et al. (2006), Hollien and Harnsberger (2006), and Brown, Senter, and Ryan (2003).

Nunes (2010) conclude that the embarrassment levels produced by a different version of the LVA software are able to discriminate between subjects that were asked to describe embarrassing products and those that were asked to describe benign and nonembarrassing products.

While the early evidence suggests that LVA may not offer meaningful emotion metrics, more recent evidence is consistent with LVA capturing meaningful markers of emotion. However, of the two LVA-based metrics that we use to proxy for positive and negative affects, the literature offers construct validity only for Cognition Level (see Hobson et al. (2011)). We therefore caution the reader that our tests are ultimately joint tests of the hypothesis that market participations react to managerial affective states and that we are capturing affective states through the measures generated by the LVA software. Our analysis in Section V.A provides reassuring evidence with respect to this latter point, as we do observe systematic associations between our LVA measures and standard acoustic features from the vocal waveform commonly studied in the emotion literature (Owren and Bachorowski (2007)).

III. Sample Selection, Variable Measurement, and Descriptive Statistics

We derive our sample of audio files from all conference calls held between January 1 and December 31, 2007 available on the Thomson Reuters StreetEvents database. We face two main challenges with processing the audio files available on this database. First, Thomson Reuters does not retain audio files indefinitely. Rather, it archives the audio files for a time period ranging from 90 days to 1 year following the conference call date, after which they are no longer available to database subscribers.[12] Second, StreetEvents provides access to audio files as playback only, thus the audio files cannot be downloaded directly. Together, these issues impose a time constraint on our analysis of the audio files, as we must manually play and analyze the audio files while such files are available. To accommodate this constraint, we construct our sample in two phases.

In the first phase, between January 1 and March 31, 2007 we identify 2,650 conference calls for fiscal year 2006 fourth-quarter earnings where company identifiers are available on the CRSP, Compustat, and I/B/E/S databases. We remove 1,569 observations for which Thomson Reuters does not index the audio file. Audio indexing is required for meaningful voice analysis, as discussed further below. We next remove 466 observations for which the absence of data on CRSP, Compustat, or I/B/E/S prevents the construction of variables needed for the empirical tests that we employ. Thus, the final initial sample in phase I consists of 615 firm conference call observations.

---

[12] Discussions with Thomson Reuters suggest that the archiving period is primarily determined by the firms. We do not believe the choice of archiving period made by firms causes any particular self-selection bias because, in perfect foresight, all audio files could have been independently recorded from public sources and parsed apart without using Thomson Reuters StreetEvents. That is, in our setting, the use of a data provider simply reduces processing costs.

To construct our measures of managerial affect during conference calls, we play back the entire conference call audio files through LVA. The software requires a calibration period over which "normal" voice characteristics of the speaker are measured. Subsequent to calibration, LVA analyzes audio output at constant intervals relative to the calibration benchmark and produces various measures, including our variables of interest, Cognition Level and Emotion Level, which serve as the basis for negative and positive affects. LVA measurement continues until the researcher manually ends the test.

The earnings conference call audio files are uniquely suited for LVA analysis for three reasons. First, firm executives commonly begin the conference call with mundane introductions of the conference call participants and Safe Harbor statements. These "boilerplate" opening statements are ideal for calibrating the voice of each executive because they require little cognitive investment. Second, StreetEvents uses a proprietary technology called "indexed audio" that maps audio files onto the conference call transcripts. With indexed audio, a researcher can point and click to specific locations of the conference call where a given executive speaks. Since voice analysis is speaker dependent, the use of audio indexing allows us to seamlessly isolate the vocal content for a given executive throughout a conference call dialog without the confounding effects of other speakers. Finally, the LVA software is geared specifically toward settings in which subjects encounter intense interrogation, and hence we anticipate that the software is most powerful in detecting emotional states during analyst questioning.

For each conference call, we separately measure positive and negative affects for the CEO and CFO because each individual speaker has a different vocal profile that requires separate calibration. We calibrate each executive based on his introductory remarks in the call presentation. If an executive does not provide introductory remarks in the conference call, we calibrate his vocal profile using the opening moments of his speech during the conference call. The calibration is done internally in the software, and typically takes around 10 seconds to complete. We aggregate the affect measures obtained for both executives present in a call to obtain firm-level *NAFF* and *PAFF* measures.[13] LVA measures each parameter approximately 35 times per minute, implying that, for a 10-minute CEO speech, LVA will generate 350 parameter readings.

To generate conference call–level measures of *NAFF* and *PAFF*, we measure how many individual Cognition Level and Emotion Level readings from each executive were above the "critical" level as defined by the developers of LVA. We count the number of critical instances and scale it by the total number of individual readings.[14] With respect to Cognition Level, readings above 120 are indicative of severe cognitive dissonance by the subject. Hence, we use

---

[13] We do not analyze CEO and CFO affective states separately because we expect both executives to have similar information sets and similar appraisals of such information, yielding similar affective states. The Pearson correlation coefficients between CEO and CFO *PAFF* and *NAFF* measures are positive and statistically significant ($\rho = 0.28$ and 0.59 respectively; $p = 0.00$).

[14] Ex-Sense Pro-R only graphically produces the individual parameters that are needed for our empirical measures. We thank Nemesysco for accommodating our request to build a module into

the proportion of readings that have cognition levels above 120 to construct the *NAFF* measure. For *PAFF*, we measure the proportion of readings with emotion levels greater than the critical 110 level.

Panel A of Table I presents descriptive statistics for the emotion measures of the conference calls in the initial sample. The mean *PAFF* is 0.1028, indicating that, on average, managers exhibit positive affect 10% of the time during a conference call. In contrast, managers exhibit negative affect about 17% of the time (mean *NAFF* = 0.1663).

A disadvantage of a small sample from a single calendar quarter is the difficulty in drawing clear and generalizable inferences due to lack of statistical power. At the same time, the enormous costs of manual playback and analysis of individual executives throughout an entire conference call present a significant challenge, particularly because of the finite availability of the audio files. As a compromise, we expand our sample by analyzing conference call audio files of a shorter duration for the three subsequent calendar quarters of 2007.

Conceptually, the software was developed to capture the emotional states during interrogation settings in which the subject is asked questions to determine whether the subject exhibits a cognitive or emotional state different from the subject's "normal" state. Furthermore, affective states are most powerfully elicited when external stimuli are the strongest. Thus, we believe that focusing on the Q&A portion of the call gives us the best chance of success in capturing affective states.

To determine the most cost-effective duration, we partition the presentation and the Q&A portion of the initial sample of conference calls pertaining to the CEO into quartiles. We focus on the CEO rather than the CFO because the CEO arguably has the most knowledge about, and is most responsible for, a firm's performance. Moreover, CEOs tend to speak more during conference calls relative to CFOs (Li et al. (2009)). In our initial sample, we find that the average number of words spoken by the CEO (3,186) is statistically and economically greater than the average number of words spoken by the CFO (1,928).

We analyze the distribution of the two measures *PAFF* and *NAFF* for the CEO as the call progresses so as to identify the particular portion of the conference call that would be both economically and statistically meaningful. Results presented in Panels B and C of Table I suggest that both emotion measures display a gradually increasing trend throughout the conference call, consistent with what one would expect as a speaker approaches and begins to answer questions from an analyst audience in real time. In addition, we find a pronounced increase in *NAFF* during the first quartile of the Q&A portion of the call (average CEO *NAFF* increases by 6.10%, from 16.94 to 17.97).

On the basis of conceptual underpinnings and the preceding analysis, we augment our initial sample by collecting the first 5 minutes of the CEO

---

the software that allows us to extract the numerical values of the two vocal attributes we study, which are otherwise available only in graphical format. See the Internet Appendix for a screen shot of this graphical format.

Table I

Descriptive Statistics on Affective State Variables for the Initial Sample

This table reports descriptive statistics on the affective state variables calculated for an initial sample of 615 fiscal year 2006 fourth-quarter earnings conference calls occurring between January 1 and March 31, 2007. In Panel A, *PAFF* is positive affect measured for both CEO and CFO during the entire conference call; *NAFF* is negative affect measured for both CEO and CFO during the entire conference call. Panel B reports how *PAFF* evolves over the course of the conference call for CEOs. *PAFF* is calculated as in Panel A, except that it is only calculated for the CEO, and is measured at eight intervals: the four quintiles of the presentaiton portion of the conference call and the four quintiles of the Q&A session. Panel C reports how *NAFF* evolves over the course of the conference call for CEOs. *NAFF* is calculated as in Panel A, except that it is calculated only for the CEO, and is measured at eight intervals: the four quintiles of the presentation portion of the conference call and the four quintiles of the Q&A session. See Appendix A for a detailed description of *PAFF* and *NAFF*.

Panel A: Descriptive Statistics of *PAFF* and *NAFF*					
Variable	N	Mean	Std. Dev.	Median	Min	Max
PAFF	615	0.1028	0.0174	0.1039	0.0347	0.1610
NAFF	615	0.1663	0.0644	0.1632	0.0256	0.3372

Quartiles	Mean	Std. Dev.	Median	Mean Change	Mean Change (%)	p-Value Mean Change = 0
Panel B: Descriptive Statistics Across Sections of the Conference Call: CEO *PAFF*						
Presentation Section						
1	0.0948	0.0440	0.0934			
2	0.0978	0.0442	0.0956	0.0030	3.20%	
3	0.0972	0.0423	0.0967	−0.0006	0.60%	0.79
4	0.1061	0.0534	0.1004	0.0089	9.20%	0.00
Q&A Section						
1	0.1078	0.0402	0.1072	0.0017	1.60%	0.51
2	0.1103	0.0434	0.1083	0.0025	2.30%	0.24
3	0.1118	0.0441	0.1125	0.0015	1.40%	0.47
4	0.1132	0.0372	0.1117	0.0014	1.20%	0.51
Panel C: Descriptive Statistics Across Sections of the Conference Call: CEO *NAFF*						
Presentation Section						
1	0.1542	0.09667	0.1467			
2	0.1565	0.08869	0.1547	0.0022	1.50%	0.65
3	0.1601	0.09119	0.1538	0.0036	2.30%	0.45
4	0.1694	0.09088	0.1667	0.0092	5.80%	0.06
Q&A Section						
1	0.1797	0.08396	0.1721	0.0104	6.10%	0.03
2	0.1809	0.07788	0.1753	0.0012	0.70%	0.78
3	0.1879	0.07763	0.1851	0.0070	3.90%	0.09
4	0.1794	0.07125	0.1759	−0.0086	−4.60%	0.03

responses from the question and answer portion of the conference call. By collecting a shorter duration, we may be missing out on important affect variation, because, as shown in Table I, Panels B and C, *PAFF* and *NAFF* levels are still relatively high with considerable variance at *all* points during the conference call.[15] However, a shorter duration allows us to analyze many more firm quarters over a longer time period, which increases external and statistical conclusion validity. To examine the empirical validity of using a shorter duration, for the initial sample, we estimated the correlation between the overall *PAFF* (*NAFF*) for the entire conference call with that of the *PAFF* (*NAFF*) computed for the first 5 minutes of the CEO responses during the Q&A section and find that the correlation is quite high (ρ for *PAFF* = 0.53; *NAFF* = 0.79). This finding gives us some confidence that the LVA measures computed for a shorter duration capture statistically meaningful variation in the affective states.

Our second phase of data collection yields 1,032 firm-quarter conference calls hosted from April 1 to December 31, 2007. Together, the two phases of data collection yield a final sample 1,647 observations representing 691 unique firms. Our final sample has far fewer observations in the second calendar quarter of 2007 because, by the time we made our decision to collect more data, Thomson Reuters had purged the voice files for several of our sample firms.

We obtain stock return data from the CRSP database and www.yahoo.com as necessary. We obtain financial data from the Compustat database to the extent it is available. For financial data relating to the most recent periods, we hand-collect it from the Edgar database available at www.sec.gov. We obtain analyst expectations of earnings and earnings forecast revision data from I/B/E/S.

Descriptive statistics for the combined sample are presented in Panel A of Table II. The mean (median) for *PAFF* is 0.1086 (0.1064) whereas the mean (median) for *NAFF* is 0.1758 (0.1721). These descriptives are comparable to those obtained for the initial sample (see Panel A of Table I), suggesting that the augmented sample is quite representative. The sample firms have an average (median) quarterly return on assets (*ROA*) of 0.41% (1.04%) and assets of $7.6 ($1.2) billion. The mean (median) firm has revenues of $941 million ($213 million) and market value of equity of $5.7 billion ($1.3 billion). Thus, our sample predominantly consists of large firms. In Panel B, we provide the industry composition for our sample firms. While we do not observe significant industry clustering, the sample contains a relatively greater number of firms from the computer, financial, and services industries.

The Pearson correlation matrix of all the financial variables and the two affect measures are presented in Panel C of Table II. Several observations are worth noting. First, *NAFF* is negatively related to size (*LNMVE*) ($\rho = -0.15$, $p = 0.00$), negatively related to firm profitability (ρ (*ROA,NAFF*) = -0.09, $p = 0.00$), and positively related to volatility (ρ (*VOL,NAFF*) = 0.11,

---

[15] The presence of some emotion during the presentation portion of the conference is not surprising. Managers rationally anticipate some of the questions analysts will likely ask when preparing the presentation portion of the conference call, thereby endogenizing some the emotional effects that would otherwise be present during the Q&A period of the conference call.

Table II
Descriptive Statistics and Sample Characteristics

This table reports descriptive statistics and sample characteristics for 1,647 quarterly earnings conference calls occurring between January 1 and December 31, 2007. Panel A reports descriptive statistics for the sample observations. Panel B reports industry concentrations for the sample observations. Panel C reports correlations between positive and negative emotional states and sample firm characteristics. See the Appendix for a detailed description of the variables.

Panel A: Descriptive Statistics						
Variable	N	Mean	Std. Dev.	Median	Min	Max
PAFF	1,647	0.1086	0.0245	0.1064	0.0199	0.2391
NAFF	1,647	0.1758	0.0702	0.1721	0.0000	0.4570
ROA	1,647	0.0041	0.0455	0.0104	−0.2059	0.1156
STDROA	1,647	0.0148	0.0237	0.0063	0.0001	0.1523
ASSETS	1,647	7,658	21,441	1,227	29	143,369
NEGWORDS	1,647	0.0087	0.2243	0.0108	−1.0100	1.6000
POSWORDS	1,647	0.0092	0.2665	0.0105	−1.2000	0.9800
FREV	1,647	−0.0016	0.0056	−0.0002	−0.0344	0.0121
RECREV	1,647	0.0017	0.2006	0.0000	−2.0000	1.5000
FDISP	1,647	0.0344	0.0519	0.0200	0.0000	0.3600
CAR(0,1)	1,647	−0.0023	0.0787	−0.0019	−0.4983	0.3319
CAR(2,180)	1,647	−0.0612	0.4218	−0.0210	−3.6617	1.7641
UE_t	1,647	−0.0008	0.0129	0.0004	−0.0903	0.0310
UE_{t+1}	1,647	−0.0016	0.0177	0.0004	−0.1344	0.0409
UE_{t+2}	1,146	−0.0020	0.0210	0.0004	−0.1719	0.0383
$UE_{t+1,t+2}$	1,146	−0.0028	0.0312	0.0009	−0.3063	0.0792
LNMVE	1,647	7.2762	1.5544	7.1625	3.9519	11.4565
MOM	1,647	−0.0035	0.2404	0.0035	−0.7190	0.6579
BM	1,647	0.4404	0.2854	0.3942	−0.1132	1.4327
VOL	1,647	0.0212	0.0086	0.0198	0.0078	0.0514

Panel B: Industry Composition				
	Sample Firms		All Compustat Firms	
Industry	N	%	N	%
Chemicals	30	1.82	411	1.82
Computers	236	14.33	2,908	12.85
Extractive	59	3.58	904	3.99
Financial	218	13.24	3,050	13.48
Food	23	1.40	401	1.77
Insurance/RealEstate	117	7.10	2,306	10.19
Manf:ElectricalEqpt	51	3.10	767	3.39
Manf:Instruments	110	6.68	1,062	4.69
Manf:Machinery	28	1.70	544	2.40
Manf:Metal	20	1.21	473	2.09
Manf:Misc.	8	0.49	214	0.95
Manf:Rubber/glass/etc	9	0.55	371	1.64
Manf:TransportEqpt	30	1.82	340	1.50
Mining/Construction	28	1.70	622	2.75
Pharmaceuticals	124	7.53	900	3.98
Retail:Misc.	93	5.65	933	4.12
Retail:Restaurant	19	1.15	286	1.26

(continued)

Table II—*Continued*

Panel B: Industry Composition				
	Sample Firms		All Compustat Firms	
Industry	N	%	N	%
Retail:Wholesale	28	1.70	781	3.45
Services	178	10.81	2,064	9.12
Textiles/Print/Publish	80	4.86	845	3.73
Transportation	102	6.19	1,388	6.13
Utilities	50	3.04	658	2.91
Not assigned	6	0.36	405	1.79
Total	1,647	100.00	22,633	100.00

Panel C: Pearson Correlations among Emotion Levels and Firm Characteristics (Significance Levels in Parentheses)

	PAFF	*NAFF*
NAFF	0.04	
	(0.11)	
ROA	0.01	−0.09
	(0.80)	(0.00)
STDROA	0.01	0.04
	(0.63)	(0.15)
ASSETS	0.01	−0.12
	(0.82)	(0.00)
POSWORDS	−0.03	−0.01
	(0.20)	(0.60)
NEGWORDS	−0.01	0.03
	(0.70)	(0.21)
FREV	−0.00	−0.03
	(0.86)	(0.26)
RECREV	0.03	0.01
	(0.21)	(0.61)
FDISP	0.02	−0.00
	(0.49)	(0.91)
CAR(0,1)	0.05	−0.04
	(0.05)	(0.13)
CAR(2,180)	−0.00	−0.05
	(0.87)	(0.05)
UE_t	−0.00	−0.04
	(0.93)	(0.14)
UE_{t+1}	0.01	−0.04
	(0.66)	(0.12)
UE_{t+2}	0.02	−0.09
	(0.40)	(0.00)
$UE_{t+1,t+2}$	0.01	−0.09
	(0.68)	(0.00)
LNMVE	0.01	−0.15
	(0.56)	(0.00)
MOM	−0.01	−0.05
	(0.67)	(0.06)
BM	0.02	−0.01
	(0.32)	(0.79)
VOL	0.02	0.11
	(0.42)	(0.00)

$p = 0.00$). These correlations provide initial evidence on the construct validity for the *NAFF* variable derived from the LVA software. Recall that *NAFF* is purported to capture cognitive dissonance. For managers who believe they are competent and in control of their firms, poor accounting performance will cause cognitive dissonance because it undermines the manager's belief about competency. Additionally, if small firms and firms with high volatility capture settings that are more uncertain, it is likely that managers who believe they are in control of the firm will experience cognitive dissonance. We do not find statistically significant correlations between *PAFF* and the aforementioned variables, however.

Second, we do not observe a strong systematic relation between the two affect measures ($\rho = 0.04$, $p = 0.11$). This finding is not surprising because managers discuss many issues during a conference call, each of which may induce a positive or negative affect on the manager.[16] Furthermore, research suggests that positive and negative affect need not be negatively correlated (Diener and Emmons (1985), Cacioppo and Bernston (1994)). The lack of relation also suggests that neither affect measure subsumes the other.

Third, we find some evidence that the affect variables convey information to the capital markets. The contemporaneous market reaction to *NAFF* (*PAFF*) is weakly (significantly) negative (positive) and of similar absolute magnitude ($\rho = -0.04$, $p = 0.13$; $\rho = 0.05$, $p = 0.05$). Further, for *NAFF*, we find a negative association with earnings news two quarters in the future, UE_{t+2}, ($\rho = -0.09$, $p = 0.00$). Since earnings news is based on analyst expectations of future earnings, the association of *NAFF* with future earnings news implies that analysts have not taken into account the implications of negative affect into their earnings forecasts contemporaneously. The negative correlation between *NAFF* and stock returns over the subsequent 180 days ($\rho = -0.05$, $p = 0.05$) suggests that investors appear to incorporate the implications of *NAFF* for future earnings news with some delay.[17] Collectively, these results provide initial evidence to suggest there is information in affect conveyed via voice, and that the implications of negative affect take longer to get incorporated into price. Naturally, to draw more definitive conclusions about the role of affect as an information source and how market participants incorporate such information, we must rule out confounding factors. We do so in our multivariate tests that follow.

[16] Some firms explicitly attempt to provide a balanced view of the firm such that a portion of their conference call presentation is dedicated to positive aspects of the firm and another portion to negative aspects. Managers may provide a balanced perspective in the Q&A section as well. For example, Cisco Systems noted the following in its 2004 first-quarter earnings conference call: "Reminding those who have limited exposure to our prior conference calls, we try to give equal balance to both what went well and our concerns." Explicit balancing implies a positive correlation between *NAFF* and *PAFF*, as each unit of *NAFF* is balanced with a unit of *PAFF*.

[17] Johnson (2004) argues that firms with high idiosyncratic uncertainty have increased option values that expire over time and yield negative future stock returns. Since *NAFF* is positively correlated with idiosyncratic return volatility (*VOL*), an alternative explanation for the negative relation between *NAFF* and future stock returns is that *NAFF* simply captures firm-specific idiosyncratic uncertainty. In our multivariate analysis, we control for idiosyncratic return volatility.

IV. Results

A. *Do Market Participants Respond to Managerial Affect?*

We begin by assessing whether investors respond to managerial affect by examining the contemporaneous stock market reaction to vocal cues. We estimate daily abnormal returns using the returns on the size and book-to-market portfolio in which the firm resides (Fama and French (1993)) as the benchmark return, and then regress the 2-day cumulative abnormal returns (*CAR*s) measured around the conference call date (*CAR*(0, 1)) on the vocal cue measures, *NAFF* and *PAFF*. We control for quantitative accounting news contained in the earnings conference call by including the magnitude of unexpected earnings (*UE*), with expectations based on the last summary consensus median analyst forecast prior to the earnings conference call. We expect a positive coefficient on *UE*.

We next consider whether the vocal cue–based measures capture information incremental to that contained in the linguistic tone documented in prior research. We use the positive-word and negative-word dictionaries of Loughran and McDonald (2011) to compute the unexpected percentage of positive words (*POSWORDS*) and negative words (*NEGWORDS*) in the entire conference call dialog.[18] We expect a positive (negative) coefficient on *POSWORDS* (*NEGWORDS*).

We control for size, growth, and risk, which have been shown to be related to market returns (Collins and Kothari (1989)). As empirical proxies for size, growth, and risk, respectively, we use the natural logarithm of market value of equity at the end of the current quarter (*LNMVE*); book-to-market (*BM*), calculated as the book value of shareholders equity at the end of the current quarter scaled by the market value of equity; and return volatility (*VOL*), measured as the standard deviation of daily stock returns over the 125 trading days prior to the earnings announcement. Finally, we control for return momentum (*MOM*), measured as the cumulative daily abnormal return over the 125-day trading window [–127, –2] prior to the earnings announcement. We estimate the following specification:

$$CAR(0, 1) = \lambda_0 + \lambda_1 PAFF + \lambda_2 NAFF + \lambda_3 UE_t + \lambda_4 LNMVE + \lambda_5 MOM + \lambda_6 BM$$
$$+ \lambda_7 VOL + \lambda_8 POSWORDS + \lambda_9 NEGWORDS + \varphi. \quad (1)$$

We estimate equation (1) using pooled ordinary least squares regression with robust standard errors.[19] Column (1) of Table III presents the results of estimating equation (1). As expected, the coefficient on unexpected earnings (UE_t) is positive and statistically significant, suggesting that the market responds to earnings news. Consistent with Davis et al.'s (2010) analysis of earnings press releases, we find a statistically significant positive (negative) relation between *POSWORDS* (*NEGWORDS*) and contemporaneous returns.

---

[18] The positive word and negative word dictionaries are available at http://www.nd.edu/∼mcdonald/Word_Lists.html.

[19] All our empirical results are robust to clustering standard errors by firm.

Table III
Estimation of the Association between Affect and Contemporaneous Stock Returns

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and the contemporaneous stock market reaction (*CAR*(0,1)). Robust standard errors are presented in parentheses below the coefficient estimates. *** and **: significant at 0.01 and 0.05 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	(1)	(2)
Intercept	?	−0.0108	−0.0034
		(0.0205)	(0.0200)
PAFF	+	0.1647**	
		(0.0776)	
NAFF	−	−0.0290	
		(0.0270)	
$PAFF^{HS}$	+		0.1263*
			(0.0961)
$NAFF^{HS}$	−		−0.1522***
			(0.0440)
$PAFF^{LS}$	+		0.1507**
			(0.0817)
$NAFF^{LS}$	−		0.0432
			(0.0316)
UE_t	+	0.8204***	0.2576
		(0.2494)	(0.2681)
LNMVE	?	0.0003	−0.0008
		(0.0015)	(0.0015)
MOM	?	0.0040	0.0011
		(0.0107)	(0.0104)
BM	?	−0.0037	−0.0006
		(0.0073)	(0.0071)
VOL	?	−0.1926	−0.2094
		(0.3432)	(0.3310)
POSWORDS	+	0.0290***	0.0236***
		(0.0072)	(0.0071)
NEGWORDS	−	−0.0453***	−0.0399***
		(0.0086)	(0.0085)
N		1,647	1,647
Adjusted R^2		7.64%	10.65%

More important, with respect to our variables of interest, we observe a significantly positive relation between positive affect (*PAFF*) and returns (coefficient = 0.1647; $p < 0.05$). However, the coefficient on negative affect (*NAFF*), although negative, does not achieve statistical significance at conventional levels. This result indicates that, on average, investors perceive positive information from positive affect but no information from negative affect. There are two possible explanations for the weak result for *NAFF*. Investors may be optimistic, on average, and fail to incorporate the negative affective state in comparison to the positive affective state. An alternative explanation is that, during the conference call, the analysts' questions are not scrutinizing enough

to trigger a negative affective state. To test these competing explanations, we identify situations in which the analysts are most likely to scrutinize and interrogate managers during conference calls.

Recent survey evidence by Graham et al. (2005, p. 42) points to such a situation: "CFOs dislike the prospect of coming up short on their numbers, particularly if they are guided numbers, in part because the firm has to deal with extensive interrogations from analysts about the reasons for the forecast error, which limits their opportunity to talk about long-run strategic issues." Accordingly, we posit that managers of firms who miss analysts' earnings benchmarks are most likely to be extensively interrogated, in turn evoking affective states.[20]

To test this hypothesis, we define high-scrutiny affect, $PAFF^{HS}$ ($NAFF^{HS}$), as $PAFF$ ($NAFF$) when UE_t is less than zero, and zero otherwise. We define low-scrutiny affect, $PAFF^{LS}$ ($NAFF^{LS}$), as $PAFF$ ($NAFF$) when UE_t is greater than or equal to zero, and zero otherwise. These definitions allow the coefficient on $PAFF$ and $NAFF$ to vary depending on whether the firm is in a high- or low-scrutiny setting, where scrutiny is based on sign of the deviation of the firm's reported earnings from analyst expectations. We then estimate the following specification:

$$
\begin{aligned}
CAR(0, 1) = {} & \lambda_0 + \lambda_1 PAFF^{HS} + \lambda_2 NAFF^{HS} + \lambda_3 PAFF^{LS} + \lambda_4 NAFF^{LS} + \lambda_5 UE_t \\
& + \lambda_6 LNMVE + \lambda_7 MOM + \lambda_8 BM + \lambda_9 VOL + \lambda_{10} POSWORDS \\
& + \lambda_{11} NEGWORDS + \varphi.
\end{aligned}
\tag{2}
$$

Regression results from estimating equation (2) are presented in column (2) of Table III. Allowing the effects of $PAFF$ and $NAFF$ to vary by the extent of scrutiny improves the model fit substantially, as evidenced by the increase in adjusted R^2 from 7.64% to 10.65%. Our evidence is consistent with the idea that analysts offer greater scrutiny when earnings expectations are not met and that this scrutiny in turn evokes emotions, rather than constituting a failure on the part of investors to incorporate negative affective states exhibited by CEOs.[21] In particular, we find that the coefficient on $NAFF^{HS}$ is negative

---

[20] Empirical analysis is consistent with more analyst scrutiny in conference calls when firms miss analyst expectations. We use all transcripts in Thomson StreetEvents during the period 2002 to 2004 and examine the words spoken between management and each individual analyst during the Q&A session. We find that the firms that miss analyst forecasted earnings have less positive (more negative) dialogs during the Q&A. Consistent with increased scrutiny, even analysts with relatively favorable stock recommendations, who otherwise exchange more favorable words with management, are more negative when the firm misses earnings targets.

[21] We find that the coefficient on unexpected earnings is no longer statistically significant after allowing the effects of emotion to vary in high- and low-scrutiny conditions. This finding should not be interpreted as information in emotion subsuming the effects of quantitative earnings news. The relation between contemporaneous stock returns and unexpected earnings has been shown to be nonlinear (Freeman and Tse (1992)). Accommodating this nonlinearity in the earnings–returns relation reveals that stock returns are increasing in unexpected earnings in a statistically significant way and the inferences on our emotion-based variables remain unchanged.

and statistically significant (coefficient -0.1522; *p*-value < 0.01). The coefficient on $PAFF^{HS}$ is positive and statistically significant (coefficient 0.1263; *p*-value < 0.10), and is of similar magnitude as that on $NAFF^{HS}$. While there are no observed differences in the market perceptions of positive affective state across scrutiny conditions (an F-Test for the equality of the coefficients on $PAFF^{HS}$ and $PAFF^{LS}$ cannot be rejected), the market does not react to negative affect for firms in low-scrutiny conditions ($NAFF^{LS} = 0.0432$, *p*-value > 0.10) and an F-Test for the equality of the coefficients on $NAFF^{HS}$ and $NAFF^{LS}$ is rejected (*p*-value < 0.01). These results imply that the market reaction to negative affect is statistically greater, and in fact only exists when firms are in high-scrutiny conditions.[22]

To better understand the nature of the information contained in managerial affect, we examine whether and how analysts incorporate the signals in the vocal cue–based measures when revising their expectations about a firm's financial future. If the news in positive (negative) affect is informative to analysts, we would expect to see upward (downward) revisions in earnings forecasts, stock recommendations, or both. Given the results in Table III that investors react to both positive and negative affect in high-scrutiny conditions, we will pay particular attention to that setting in our remaining empirical analysis.

To examine analyst reactions, we use analyst forecast revisions (*FREV*) and changes in analyst recommendations (*RECREV*) as the dependent variables instead of $CAR(0, 1)$ in equation (2). We include the contemporaneous market reaction $CAR(0, 1)$ as an additional explanatory variable to control for news in the earnings release that is not quantifiable by other variables in the model. We expect a positive coefficient on $CAR(0, 1)$ consistent with the prediction for other news proxies. Our expectations for the other explanatory variables are identical to those for equation (2).

We measure analyst forecast revisions (*FREV*) as the one-quarter-ahead forecast revision representing the difference between the median one-quarter-ahead forecast issued after and before the current-period earnings announcement, scaled by stock price 2 days preceding the conference call.[23] The median forecast before (after) the current period earnings announcement is determined using the last (first) forecast of all individual analysts issuing forecasts during the 90-day period before (after) the current quarter earnings announcement date. We measure recommendation revision (*RECREV*) as the difference between the average consensus recommendation immediately after and before the earnings announcement. Consensus recommendations are

---

[22] A competing explanation for this finding is that investors are more attuned to the conversation during conference calls when earnings expectations are missed, rather than analysts providing more extensive scrutiny. However, our subsequent finding that investors fail to fully incorporate the implications of negative affect when earnings expectations are not met is inconsistent with this explanation (see Table VI).

[23] Observations (128 firms) with no individual analyst forecast revisions during the period are set to zero values. We reestimated the regression after eliminating the 128 firms and our inferences remain unchanged.

Estimation of the Association between Affect and Analyst One-Quarter-Ahead Forecast Revisions

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and analyst earnings forecast revisions (*FREV*) and recommendation revisions (*RECREV*). See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	*FREV* (1)	*RECREV* (2)
Intercept	?	0.0022*	−0.2723***
		(0.0013)	(0.0568)
PAFF^{HS}	+	−0.0047	0.4200**
		(0.0076)	(0.2266)
NAFF^{HS}	−	0.0019	−0.0121
		(0.0036)	(0.1025)
PAFF^{LS}	+	0.0014	0.1056
		(0.0051)	(0.1861)
NAFF^{LS}	−	−0.0002	0.0862
		(0.0018)	(0.0732)
UE$_t$	+	0.1273***	0.9383**
		(0.0256)	(0.4967)
CAR(0,1)	+	0.0111***	0.5224***
		(0.0021)	(0.0852)
LNMVE	?	−0.0001	−0.0016
		(0.0000)	(0.0035)
MOM	?	0.0031***	0.0715***
		(0.0007)	(0.0223)
BM	?	−0.0032***	−0.0014
		(0.0007)	(0.0212)
VOL	?	−0.0834***	0.2570
		(0.0247)	(0.6596)
LAGREC	−		−0.1073***
			(0.0126)
POSWORDS	+	0.0006	−0.0148
		(0.0005)	(0.0218)
NEGWORDS	−	−0.0005	−0.0194
		(0.0007)	(0.0271)
N		1,647	1,647
Adjusted R^2		25.72%	12.06%

measured as the average of recommendations across all analysts. In determining the average, strong buy recommendations are coded as 5, buy recommendations as 4, hold as 3, sells as 2, and strong sells as 1.

The results are reported in Table IV. For the forecast revision regression, the coefficients on the managerial affect measures in the high-scrutiny condition, *PAFF^{HS}* and *NAFF^{HS}*, are statistically insignificant. This implies that analysts do not take into account the information contained in the affect measures when revising their earnings expectations for the upcoming quarter. Consistent with prior work, we find that analysts significantly revise their

next-period forecast based on the nature of unexpected earnings (coefficient on $UE_t = 0.1273$; *p*-value < 0.01) and the market's interpretation of earnings (coefficient on $CAR(0,1) = 0.0111$; *p*-value < 0.01).[24] We find no association between linguistic tone and forecast revision activity.

One plausible explanation for our finding that analysts do not incorporate managerial affect is that the news does not map into a firm's near-term earnings. Rather, it contains soft information that analysts incorporate into their longer-term projections of firm performance in metrics such as stock recommendations (Bradshaw (2004)). To consider this explanation, we investigate changes in analyst recommendations, arguably a broader measure and admittedly a coarser measure than changes in analyst expectations. We include the level of analyst recommendations immediately prior to the earnings announcement (*LAGREC*) as an additional independent variable to control for potential nonlinearity in the changes variables given the truncated distribution of the level of recommendations. For example, an increase (decrease) in recommendation cannot occur for a recommendation that is already a strong buy (sell). Including *LAGREC* also controls for potential mean reversion in recommendations. We predict a negative coefficient for *LAGREC*.

Results presented in column (2) of Table IV indicate that, in the high-scrutiny condition, analysts on average incorporate the positive affective state when making recommendation changes (coefficient of $PAFF^{HS} = 0.4200$; *p*-value < 0.05). We do not find significant results for negative affective state under either scrutiny condition, which is consistent with two potential explanations. Either analysts do not understand negative affect, or analysts do understand negative affect and act on incentives to avoid incorporating such negative information into their stock recommendations (O'Brien et al. (2005)). Subsequent analysis provides more support for the latter explanation.

We do not observe a relation between linguistic measures and recommendation changes. Combining this result with the finding in Engelberg (2008) that linguistic tone in earnings press releases predicts future stock returns, one can conclude that part of the reason why linguistic tone predicts future returns is that analysts do not alert investors to the implications of linguistic tone for future performance.

To summarize, in high-scrutiny settings in which the ability to detect emotional states is most pronounced, the contemporaneous reactions by investors provide support for the hypothesis that investors perceive news in vocal cue measures. Analysts do not appear to incorporate the information contained in vocal cue measures in determining near-term earnings forecasts, but do so asymmetrically in their stock recommendations.

---

[24] Inclusion of the contemporaneous market reaction as a proxy for other earnings information that we cannot explicitly control for may result in controlling away the potential effects of *NAFF* and *PAFF*. Reestimating Table IV, column (1), by excluding $CAR(0,1)$ from the estimation yields inferences that are similar to those presented.

B. Does Managerial Affect Predict Future Firm Performance?

In this section, we formally investigate whether vocal cues provide insights into managerial affective states that in turn are informative about future firm performance. In particular, we test the hypothesis that investors' response to vocal cues is consistent with the idea that these measures provide novel information about future earnings realizations. We focus on unexpected future earnings as a proxy for the potential cash flow news inherent in the capital market response. Specifically, we use future analyst forecast error scaled by stock price 2 days before the period t conference call (*UE*) as our proxy for future firm performance. Analyst forecast error is computed as the difference between actual earnings per share minus the summary consensus median earnings forecast immediately prior to the earnings announcement.

We estimate the following empirical specification:

$$UE_{t+1,t+2} = \beta_0 + \beta_1 PAFF^{HS} + \beta_2 NAFF^{HS} + \beta_3 PAFF^{LS} + \beta_4 NAFF^{LS} + \beta_5 UE_t$$
$$+ \beta_6 FREV + \beta_7 FDISP + \beta_8 LNMVE + \beta_9 MOM + \beta_{10} BM + \beta_{11} VOL$$
$$+ \beta_{12} POSWORDS + \beta_{13} NEGWORDS + \varepsilon. \tag{3}$$

We consider unexpected earnings up to two quarters ahead because we are unable to obtain future earnings for all the firms beyond two quarters.[25] If the vocal cue measures contain useful information about future performance consistent with market perception, we should expect a positive (negative) co-efficient on $PAFF^{HS}$ ($NAFF^{HS}$). In equation (3), we also include several other control variables that have been shown to affect future unexpected earnings such as analyst forecast revisions, forecast dispersion, firm size, return momentum, book-to-market ratio, and *VOL*. We measure forecast dispersion (*FDISP*) as the standard deviation of analysts' earnings per share forecasts derived from the distribution of I/B/E/S consensus earnings per share forecasts immediately prior to the earnings announcement.[26] All other variables are as previously defined.

Table V presents the regression results from estimating equation (3). We present results from one-period-ahead and two-period-ahead unexpected earnings in column (1) and column (2), respectively. In column (3), we report results using aggregate unexpected earnings for the two periods as the dependent variable. The coefficient on unexpected earnings is positive and statistically significant across all three columns, indicating persistence in unexpected earnings. Results in column (1) indicate that, although higher levels of excitement (*PAFF*) and cognitive dissonance (*NAFF*) exhibited by executives under both scrutiny conditions are positively and negatively associated with future unexpected earnings, the statistical significance does not reach acceptable levels. This may be because the information contained in the vocal cues extends beyond one period. Consistent with this conjecture, we find that the vocal

---

[25] Our results for three-quarters-ahead unexpected earnings for a reduced sample are similar.

[26] Inferences are unchanged when we scale the dispersion of analyst forecasts with either the absolute value of actual earnings per share or the standard deviation of earnings per share.

Table V
Estimations of the Association between Affect and Future Earnings News

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and future earnings surprises (*UE*). Superscripts *HS* and *LS* represent high- and low-scrutiny partitions. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	UE_{t+1} (1)	UE_{t+2} (2)	$UE_{t+1,t+2}$ (3)
Intercept	?	0.0081	0.0157**	0.0224**
		(0.0045)	(0.0068)	(0.0106)
$PAFF^{HS}$	+	0.0234	0.0690**	0.0753*
		(0.0231)	(0.0338)	(0.0510)
$NAFF^{HS}$	−	−0.0027	−0.0307**	−0.0431*
		(0.0129)	(0.0185)	(0.0294)
$PAFF^{LS}$	+	0.0107	0.0207	0.0215
		(0.0151)	(0.0223)	(0.0291)
$NAFF^{LS}$	−	−0.0056	−0.0121	−0.0192*
		(0.0053)	(0.0114)	(0.0118)
UE_t	+	0.4767***	0.4424***	0.7472***
		(0.1104)	(0.1531)	(0.2429)
FREV	+	0.4298**	0.2856	0.5811*
		(0.1878)	(0.2663)	(0.3921)
FDISP	−	−0.0357**	−0.0447*	−0.0824*
		(0.0163)	(0.0295)	(0.0521)
LNMVE	?	−0.0003	−0.0006	−0.0007
		(0.0003)	(0.0005)	(0.0008)
MOM	?	0.0078***	0.0092***	0.0123***
		(0.0028)	(0.0031)	(0.0047)
BM	?	−0.0074***	−0.0144***	−0.0211***
		(0.0023)	(0.0039)	(0.0061)
VOL	?	−0.1355	−0.3020**	−0.3334*
		(0.0977)	(0.1465)	(0.1949)
POSWORDS	+	−0.0020	−0.0022	−0.0028
		(0.0015)	(0.0018)	(0.0026)
NEGWORDS	−	0.0013	−0.0030	−0.0022
		(0.0021)	(0.0038)	(0.0053)
N		1,647	1,146	1,146
Adjusted R^2		28.80%	17.12%	20.25%

cue measures in the high-scrutiny condition predict two-period-ahead unexpected earnings (coefficient on $PAFF^{HS}$ is 0.0690, *p*-value < 0.05; coefficient on $NAFF^{HS}$ is −0.0307, *p*-value < 0.05). Our findings are similar when we combine one-period-ahead and two-period-ahead unexpected earnings (column (3)).

These findings hold after controlling for information contained in the reported earnings number and the linguistic content of the conference calls. Not surprisingly, unexpected earnings are a potent predictor of future unexpected earnings. We find no statistical association, however, between words spoken during the conference call and future unexpected earnings. Overall, our

findings suggest that affective states possess incremental information to linguistic tone in predicting future unexpected earnings, particularly when managers miss analysts' current-period earnings estimates.

C. *Does Managerial Affect Predict Future Stock Returns?*

Based on the evidence thus far, we can conclude that nonverbal vocal cues have significant information content for future firm performance, particularly when analysts scrutinize managers' statements. However, the evidence presented with respect to analyst forecast revisions and recommendation changes suggests that analysts may not fully incorporate the information contained in these vocal cues. Past accounting literature documents price drift with respect to quantitative earnings information (Bernard and Thomas (1989)) and with respect to optimistic and pessimistic language (Demers and Vega (2010), Engelberg (2008)). To the extent that the market fails to fully appreciate the implications of the nonverbal signals we investigate here, we expect a systematic relation between vocal cues and future stock returns. Therefore, in this section, we test whether this information subsequently becomes incorporated into stock price.

Alternatively stated, we examine whether the affect measures predict future abnormal returns. Because stock returns reflect revisions in expectations about future cash flows and earnings, we expect that the information contained in the affect measures, although not fully incorporated in contemporaneous market returns, will be incorporated in future stock returns when the implications of these measures for future fundamentals are subsequently realized.

We test this prediction by using the *CAR*s over the 180 trading days following 2 days after the earnings conference call (*CAR*(2,180)). We restrict our analysis to 180 days because we hand-collect returns data from www.yahoo.com and we stop data collection in September 2008. Recall from the correlations in Table II that we observe a negative statistical association between negative affect (*NAFF*) and future abnormal returns (*CAR*(2,180)) whereas there is no statistically significant correlation between *PAFF* and future abnormal returns. The univariate results, however, do not account for other factors known to predict long-run returns. We, therefore, estimate the following multivariate model that controls for risk and other factors shown to determine future stock returns:

$$CAR(2,180) = \omega_0 + \omega_1 PAFF^{HS} + \omega_2 NAFF^{HS} + \omega_3 PAFF^{LS} + \omega_4 NAFF^{LS} + \omega_5 UE$$
$$+ \omega_6 LNMVE + \omega_7 MOM + \omega_8 VOL + \omega_9 BM + \omega_{10} POSWORDS$$
$$+ \omega_{11} NEGWORDS + \xi. \tag{4}$$

The independent variables in model (4) are identical to those in model (2). Results of estimating equation (4) are presented in Table VI. The coefficient on UE_t is positive but not statistically significant (coefficient = 2.2959). This statistical insignificance is not surprising given that our sample comprises large firms and prior research shows that post-earnings announcement drift

Table VI

OLS Estimations of the Association between Affect Variables and Future Stock Returns

This table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and future stock returns (*CAR*(2,180)). Superscripts *HS* and *LS* represent high- and low-scrutiny partitions. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	
Intercept	?	0.1900
		(0.1123)
$PAFF^{HS}$	+	0.5280
		(0.5649)
$NAFF^{HS}$	−	−0.6463***
		(0.2580)
$PAFF^{LS}$	+	−0.2258
		(0.4790)
$NAFF^{LS}$	−	0.0104
		(0.1697)
UE_t	+	2.2959
		(1.8075)
LNMVE	?	−0.0127
		(0.0085)
MOM	?	0.2404***
		(0.0553)
BM	?	−0.0974**
		(0.0487)
VOL	?	−3.7099*
		(1.9100)
POSWORDS	+	0.0226
		(0.0397)
NEGWORDS	−	0.0161
		(0.0476)
N		1,647
Adjusted R^2		5.39%

is much less pronounced in large firms (Bernard and Thomas (1989)). We find no statistical association between linguistic tone and future abnormal returns, consistent with Loughran and McDonald (2011), who find no predictive ability for linguistic tone and 1-year-ahead returns.[27]

Pertinent to this study, we find that the coefficient on $NAFF^{HS}$ is negative and statistically significant at the 1% level (coefficient = −0.6463). This evidence is consistent with market participants underreacting to the information in negative affect exhibited by CEOs when analysts engage in extensive scrutiny

---

[27] Other researchers (Engelberg (2008), Demers and Vega (2010)) document associations between future returns and linguistic tone when using the General Inquirer's *Harvard Psycho-Sociological Dictionary*. Using that dictionary, we also find that negative (positive) tone is negatively (positively) associated with future stock returns.

of firm managers' statements. We do not find a statistically significant association between positive affect in the high-scrutiny condition and future returns (coefficient on $PAFF^{HS} = 0.5280$). For the low-scrutiny condition, neither $NAFF$ nor $PAFF$ is statistically significant.

It is difficult to ascertain why market participants fail to fully incorporate negative affect. One plausible explanation is consistent with our findings relating to asymmetric analyst recommendation changes following the conference call. Recall that in Table IV we find evidence that analysts incorporate positive affect but not negative affect into their recommendations. If part of the contemporaneous market price response is a reaction to analyst recommendation changes, the lack of downgrading for negative affect would imply a less-than-complete market reaction to negative affect. Regardless, we caution the reader that this does not establish a profitable trading strategy for two reasons. First, the time period we consider is not long enough for us to analyze calendar time returns that would help us make more definitive conclusions regarding abnormal returns. Second, any returns to a trading strategy may not be profitable after transactions costs are considered.

V. Additional Analyses

Thus far, we have three main findings: 1) vocal cues that reflect managerial affective states predict future unexpected earnings when analysts are scrutinizing managers, 2) investors and analysts respond to information contained in vocal cues at least partially, and 3) negative managerial affect predicts future returns. The inferences that we can draw from these findings depend critically on the validity of the nonverbal affect measures generated by the LVA software. In this section, we offer preliminary evidence on the construct validity of the LVA-based measures used in this study. In addition, we explore whether the predictability of future returns stems from firm-specific news releases subsequent to the conference call and from the lack of analyst activity. Finally, we conduct robustness tests to examine whether managerial affect represents information distinct from managers' innate attributes.

A. Reverse Engineering the LVA Black Box

As mentioned earlier, for proprietary reasons the LVA software developers did not provide us with specific details about how the various output measures are generated. While a comprehensive examination of the inner workings of the software is beyond the scope of this study, we investigate whether the LVA affect measures ($PAFF$ and $NAFF$) used in the study correlate with some common vocal acoustic features. In emotions research, the following acoustic source characteristics are commonly used to determine the emotion content of a speech sample (Owren and Bachorowski (2007)): mean fundamental frequency (F_0), standard deviation of fundamental frequency ($F_0(Std)$), *Jitter*, *Shimmer*, and mean harmonic-to-noise (HNR) ratio. We are not aware of a theoretical model

that offers unambiguous predictions on how each of these acoustic features is related to the LVA measures. As such, we do not posit directional predictions.

To measure these acoustic features, we use PRAAT (Boersma and Weenink (2010)), a software widely used by behavioral scientists to quantify acoustic features from digital audio files (Owren (2008)). Specifically, we stream the voice data from conference calls and encode them in mono using a sound recording program, Total Recorder 7.1 Profession Edition, and save them as uncompressed ".wav" audio files. Each audio file is then digitally analyzed using PRAAT acoustics software (version 5.2.05). We use the GSU *PRAAT* "quantifySource" add-on tool with system default settings to extract acoustic parameters (Owren (2008)).[28]

Panel A of Table VII presents results on the cross-sectional relation between the LVA measures and the five acoustic cues extracted using the PRAAT software. The results suggest that the fundamental acoustic variables are correlated with both *PAFF* and *NAFF*. Vocal perturbation measures *Shimmer* and *HNR* are associated with both *PAFF* and *NAFF*, whereas *Jitter* and the standard deviation of F_0 (F_0 (Std)) are associated with only *PAFF* and *NAFF*, respectively.[29] The explanatory power of the regression model is 33.54% for *NAFF* and 2.86% for *PAFF*. Although these associations between the LVA measures and acoustic features suggest that LVA does capture established vocal attributes, it is not obvious that the associations represent economically meaningful variation in the affective states.

To probe this further, we examine whether the association between stock returns and the LVA measures documented in Table III is due to the variation in the LVA measures stemming from the acoustic features we examine. Specifically, we take the predicted values of *PAFF* and *NAFF* using the coefficient estimates reported in Panel A of Table VII and then examine the association between these predicted values and contemporaneous stock returns (see Panel B of Table VII). The coefficient on *Predicted-NAFF* is predictably negative and particularly so in the high-scrutiny condition (see column (2)). The coefficient on *Predicted-PAFF*, however, is not statistically significant in both columns (1) and (2). This result is not surprising given the poor explanatory power of the acoustic variables for *PAFF*.

Collectively, these results suggest that the LVA measures we use capture, at least to some extent, measurable and externally verifiable acoustic features. We leave for future research a more thorough analysis of the relation between these and other LVA metrics with a broader set of acoustic features, but at a minimum our exploratory analysis is inconsistent with Lacerda's (2009) assertion that the LVA technology does not extract relevant information from the speech signal.*

---

[28] Because the PRAAT software takes a considerable amount of time when processing audio files with longer duration, we use only the first 5 minutes of the audio files in our sample to obtain the acoustic measures.

[29] These associations are robust across random subsets of the data.

* A correction was made following the initial online publication of this article on January 17, 2012, changing "consistent" to "inconsistent".

Table VII

Examining the Validity of *PAFF* and *NAFF* Measures

Panel A of this table reports OLS regression estimation of the association between managerial affect (*PAFF* and *NAFF*) and acoustic characteristics. F_0 *(Mean)* and F_0 *(Std)* are the mean and standard deviation of fundamental frequency, *Jitter* is the relative average vocal perturbation, *Shimmer* is the moment-to-moment amplitude variation, and *HNR* is the mean Harmonic-to-Noise Ratio that quantifies the degree of energy/noisiness. Panel B of this table reports ordinary least squares regression estimation of the association between managerial affect (*Predicted-PAFF* and *Predicted-NAFF*) and the contemporaneous stock market reaction (*CAR*(0, 1)). *Predicted-PAFF* (*Predicted-NAFF*) represents the predicted values of *PAFF* (*NAFF*) obtained from the estimation results reported in Panel A. Superscript *LS* (*HS*) indicates Low (High) scrutiny partitions in which UE_t is less than (greater than) zero. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses below the coefficient estimates. *** and **: significant at 0.01 and 0.05 level, respectively, in a two-tailed test (one-tailed when predicted).

Panel A: The Association between *PAFF* and *NAFF* Measures with Fundamental Acoustic Attributes		
	PAFF	*NAFF*
Intercept	0.1090***	0.2029***
	(0.0079)	(0.0207)
F_0 *(Mean)*	−0.0000	−0.0002
	(0.0000)	(0.0001)
F_0 *(Std)*	0.0001***	0.0001
	(0.0000)	(0.0001)
Jitter	−0.2122	−56.582***
	(1.2156)	(3.497)
Shimmer	−0.0141***	0.1641***
	(0.0044)	(0.0110)
HNR	0.0030***	−0.0182***
	(0.0005)	(0.0013)
N	1,647	1,647
Adjusted R^2	2.86%	33.54%

Panel B: The Association between Predicted Affect and Contemporaneous Stock Returns			
	Predicted Sign	(1)	(2)
Intercept	?	0.0550	0.0652
		(0.0578)	(0.0558)
Predicted-PAFF	+	−0.3061	
		(0.4469)	
Predicted-NAFF	−	−0.1077**	
		(0.0534)	
Predicted-PAFF^{HS}	+		−0.4567
			(0.4543)
Predicted-NAFF^{HS}	−		−0.1896***
			(0.0724)
Predicted-PAFF^{LS}	+		−0.3178
			(0.4283)
Predicted-NAFF^{LS}	−		−0.0507
			(0.0622)

(continued)

Table VII—*Continued*

Panel B: The Association between Predicted Affect and Contemporaneous Stock Returns

	Predicted Sign	(1)	(2)
UE_t	+	0.8157***	0.2346
		(0.2511)	(0.2686)
LNMVE	?	0.0002	−0.0009
		(0.0015)	(0.0015)
MOM	?	0.0043	0.0008
		(0.0107)	(0.0105)
BM	?	−0.0037	−0.0002
		(0.0073)	(0.0071)
VOL	?	−0.1874	−0.2020
		(0.3448)	(0.3334)
POSWORDS	+	0.0284***	0.0221***
		(0.0072)	(0.0071)
NEGWORDS	−	−0.0453***	−0.0406***
		(0.0086)	(0.0085)
N		1,647	1,647
Adjusted R^2		6.14%	10.51%

B. *Examining the Predictive Ability of Managerial Affect for Future Returns*

The predictive ability of managerial affect for future returns may stem from information contained in future realizations of fundamentals such as earnings or from other soft information that contains value-relevant news. We capture vocal cues that are not content specific, that is, we do not attribute the positive and negative affects arising during conference calls to specific issues that are discussed during the call. Hence, focusing on future earnings realizations would limit the implications of vocal cues for future firm fundamentals. It is plausible that the information in vocal cues captures subsequent value-relevant news events that are broader than earnings releases alone. To test this prediction, we use Lexis Nexis to obtain all press releases issued by firms (wire service stories from the company's headquarters) in our initial sample during the 180-day window following 2 days after the conference call. We then code a press release as a bad news release if the abnormal stock returns surrounding the press release (window [0,1]) are negative. We use the proportion of bad news releases during the 180-day window as our measure of information following the conference call. We then estimate a two-limit Tobit model similar to equation (4) by replacing the dependent variable with the proportion of bad news to examine if the vocal cues predict future information outcomes.

Results presented in Table VIII suggest that under both high- and low-scrutiny conditions, managers who exhibit a positive affective state have a lower proportion of bad news releases subsequent to the conference call. Similarly, managers in high-scrutiny settings who exhibit a negative affective state have a higher proportion of bad news releases, but this result is not statistically significant at conventional levels (one-tailed *p*-value = 0.13). With respect to

Table VIII

Tobit Estimations of the Association between Affect Variables and Proportion of Bad News Articles in the Future

This table reports two-limit Tobit estimation of the association between managerial affect (*PAFF* and *NAFF*) and the percentage of bad news press releases issued by the firm over the 180 days following the conference call (*PCT_BN*). *PCT_BN* is the ratio of bad news press releases issued by the firm divided by the total number of press releases issued by the firm from trading day 2 to trading day 180 after the conference call. Upper (lower) Tobit limits are set at one and zero, reflecting the bounds of the percentage-based dependent variable *PCT_BN*. Press releases are obtained by searching for the date of all wire service stories on Lexis Nexis emanating from the company's headquarters. A press release is coded as bad if the abnormal stock return on the day of the article release (or the next trading day following the article date if the article was issued on a nontrading day or after hours) is negative. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

	Predicted Sign	
Intercept	?	0.5013***
		(0.0494)
$PAFF^{HS}$	−	−0.5330**
		(0.2620)
$NAFF^{HS}$	+	0.1338
		(0.1205)
$PAFF^{LS}$	−	−0.4722**
		(0.2317)
$NAFF^{LS}$	+	−0.0016
		(0.0829)
UE_t	−	0.0765
		(0.4792)
CAR(0,1)	−	−0.0289
		(0.0742)
LNMVE	?	0.0077*
		(0.0041)
MOM	?	0.0224
		(0.0221)
BM	?	−0.0095
		(0.0190)
VOL	?	−0.0415
		(0.7024)
POSWORDS	−	0.0043
		(0.0197)
NEGWORDS	+	0.0426*
		(0.0240)
N		1,304
Mean of Dependent Variable		0.506
Log Pseudolikelihood		294.34

the linguistic tone measures, we find that negative words spoken during the conference call result in a statistically greater proportion of bad news releases, but we find no statistical association between positive words and news release proportions.

Next, we investigate the role of analysts in the observed drift in stock prices. The results thus far are consistent with the notion that either analysts

understand the negative implications of negative affect but do not incorporate it into their public signals, or analysts do not immediately understand the negative affect. To investigate this issue, we examine forecast revision activity of all analysts following the firm. If analysts understand the implications and in turn remain silent, then we should observe a negative association between *NAFF* and the percentage of analysts who revise their already-outstanding annual earnings forecasts after the earnings conference call. In contrast, if analysts do not comprehend negative affect, we should observe no association between affect and revision activity. Panel A of Table IX provides a two-limit Tobit estimation of the proportion of analysts covering each firm that revised an existing estimate of upcoming annual earnings subsequent to the conference call. The results reveal that negative affect in the high-scrutiny condition is statistically negative (coefficient of –0.3152), suggesting that, when analysts observe higher levels of negative affect, they are less likely to revise their forecasts. Interestingly, we also observe a similar effect for negative linguistic tone (coefficient of –0.0736). These findings are consistent with the hypothesis that analysts are generally reluctant to revise forecasts and recommendations downwards on receiving bad news, perhaps to obtain reciprocal benefits from managers (Westphal and Clement (2008)).

Although analysts may remain silent in their public communication, they may privately communicate their views to institutional clients (Irvine, Lipson, and Puckett (2007)). If this is the case, we should observe less drift in subsamples in which institutional investment is high. In Panel B of Table IX, we partition the data and reestimate the relation between managerial affect and future stock returns across high and low levels of institutional holdings. We find that drift for negative affect is more pronounced for firms with a low level of institutional holdings (coefficient of –0.8875) relative to those with a high level of institutional holdings (coefficient of –0.5692). However, the difference is not statistically significant. As such, we cannot definitively conclude that analysts are "tipping" institutional clients privately.

C. Robustness Tests

In our last set of tests, we examine whether the managerial affective states are merely capturing managerial attributes rather than providing information that varies with context- and firm-specific circumstances. Since we cannot be sure that the software calibration completely removes innate managerial vocal attributes, or that the emotions in voice measured by the software are completely voluntary rather than somewhat controllable, we explore the possibility that managers who have more overall and firm-specific experience may have muted emotional expressions or, at the extreme, may be able to suppress their emotional state much better than younger and less experienced managers. We test this hypothesis by including CEO age (*AGE*) and CEO tenure (*TENURE*) in the empirical specifications. We find that the relation between managerial affective states and contemporaneous stock returns is unaffected by the inclusion of *AGE* and *TENURE* (see tables presented in the Internet Appendix). This suggests that managerial attributes do not explain the

Table IX

Investigation of Analyst Incorporation of Negative Affect and Its Impact on Future Stock Returns

Panel A of this table reports two-limit Tobit estimation of the association between managerial affect and analyst revision activity. The dependent variable, *PCT_REV,* is the ratio of the number of analysts with outstanding upcoming annual earnings forecasts prior to the conference call who revise their estimates after the conference call. Panel B of this table replicates the analysis in Table VI, but allows the coefficients on affect to vary with the extent of institutional holding. Superscript *HS-HighInst* (*HS-LowInst*) represents high (low) institutional ownership in the high-scrutiny setting. Superscript *LS-HighInst* (*LS-LowInst*) represents high (low) institutional ownership in the low-scrutiny setting. High (low) institutional holdings is an indicator variable set to one if the proportion of institutional investors in a firm's stock is greater (lower) than 50%, and zero otherwise. We obtain institutional ownership from 13(f) filings during the first calendar quarter of 2007 provided in the Thomson Reuters database. See the Appendix for a detailed description of the variables. Robust standard errors are presented in parentheses. ***, **, *: significant at 0.01, 0.05, and 0.10 level, respectively, in a two-tailed test (one-tailed when predicted).

Panel A: Two-Limit Tobit Estimation of the Association between Managerial Affect and Forecast Revision Activity

	Predicted Sign	
Intercept	?	0.8530***
		(0.0555)
$PAFF^{HS}$	+	0.1043
		(0.2923)
$NAFF^{HS}$	−	−0.3152***
		(0.1359)
$PAFF^{LS}$	+	−0.3551
		(0.2363)
$NAFF^{LS}$	−	−0.1170*
		(0.0883)
UE_t	+	0.1760
		(0.5359)
$CAR(0,1)$	+	−0.0935
		(0.0686)
LNMVE	?	0.0002
		(0.0043)
MOM	?	0.0402
		(0.0250)
BM	?	−0.0594***
		(0.0229)
VOL	?	−2.7148***
		(0.8353)
POSWORDS	+	−0.0061
		(0.0222)
NEGWORDS	−	−0.0736***
		(0.0297)
N		1,647
Mean of Dependent Variable		0.704
Log Pseudolikelihood		−103.91

(continued)

Table IX—*Continued*

Panel B: OLS Estimation of the Association between Managerial Affect and Future Stock Returns across Different Levels of Institutional Ownership

	Predicted Sign	
Intercept	?	0.1875*
		(0.1132)
$PAFF^{HS-HighInst}$	+	0.3322
		(0.6010)
$NAFF^{HS-HighInst}$	−	−0.5692**
		(0.3045)
$PAFF^{LS-HighInst}$	+	−0.1815
		(0.4857)
$NAFF^{LS-HighInst}$	−	0.0121
		(0.1579)
$PAFF^{HS-LowInst}$	+	1.2332*
		(0.9621)
$NAFF^{HS-LowInst}$	−	−0.8875**
		(0.4601)
$PAFF^{LS-LowInst}$	+	−0.4354
		(0.9091)
$NAFF^{LS-LowInst}$	−	0.0296
		(0.5144)
UE_t	+	2.4257*
		(1.8152)
LNMVE	?	−0.0128
		(0.0088)
MOM	?	0.2410***
		(0.0557)
BM	?	−0.0956**
		(0.0485)
VOL	?	−3.6571*
		(1.8703)
POSWORDS	+	0.0230
		(0.0396)
NEGWORDS	−	0.1875**
		(0.1132)
N		1,647
Adjusted R^2		5.51%

information content of vocal cues. Our findings with respect to future unexpected earnings are similar to those reported previously.

VI. Conclusions

To our knowledge, this study is the first to provide evidence on the role of vocal cues as a source of information about a firm's financial prospects. We posit that vocal cues from conversations with executives during earnings conference calls convey information about the executives' affective states that in turn help predict future profitability and returns. We find that higher levels

of positive (negative) affect, as operationalized via higher levels of excitement (cognitive dissonance) determined by proprietary LVA software, conveys good (bad) news about future firm performance. That is, investors respond to the information contained in positive and negative affect as evidenced by the stock returns surrounding the conference call. The effects are most pronounced when analysts scrutinize managers during conference calls, particularly when firms miss analyst earnings estimates. More positive (negative) affect predicts two-quarter-ahead future earnings. This relation holds even after we control for quantitative information and managers' word usage during conference calls. Analysts respond asymmetrically to affective states, in that we find a positive association between recommendation changes and positive affective states but no association between recommendation changes and negative affective states. We also document that stock market participants underreact to the information contained in vocal cues containing negative affect. We do not claim, however, that such underreaction represents an arbitrageable trading strategy. Such a conclusion cannot be reached without a detailed analysis of the impact of trading costs and information acquisition costs, which would require a longer time series of data.

An important implication of our paper is that information gleaned from nonverbal cues during communications between managers and shareholders may be quite useful in resource allocation and portfolio decisions. This paper adds to the body of research in social psychology that finds an incrementally important role for nonverbal cues in communication. Future research could extend this line of inquiry in various ways. First, identifying which particular business transactions or events (such as restructurings, new customer agreements, restatements, etc.) elicit positive or negative managerial affect can potentially lead to more powerful tests and further our understanding about how vocal cues can inform investors in a capital market setting. Second, performing our analysis in other settings, like depositions and communications by the Federal Reserve Chairman, might assist in forecasting interest rates. Such an analysis would complement current work analyzing the predictive ability of the Federal Reserve Chairman's linguistic style (Piger (2006), Bligh and Hess (2007)). Finally, technological advances have increased the availability of video in addition to audio. Exploring facial expressions as yet another channel of nonverbal managerial communication in the context of financial markets would be a fruitful avenue for future research.

Appendix: Variable Definitions

Variable Name	Definition
PAFF	Positive affect measured as the percentage of spoken audio by management during the conference call with Emotion Level scores above the critical value of 110 as measured by LVA.
NAFF	Negative affect measured as the percentage of spoken audio by management during the conference call with Cognition Level scores above the critical value of 120 as measured by LVA.

(continued)

Appendix—*Continued*

Variable Name	Definition
$PAFF^{HS}$ ($NAFF^{HS}$)	Represents positive (negative) affect under a high-scrutiny partition. That is, $PAFF^{HS}$ ($NAFF^{HS}$) is set to $PAFF$ ($NAFF$) when UE_t is less than zero, and zero otherwise.
$PAFF^{LS}$ ($NAFF^{LS}$)	Represents positive (negative) affect under a low-scrutiny partition. That is, $PAFF^{LS}$ ($NAFF^{LS}$) is set to $PAFF$ ($NAFF$) when UE_t is greater than or equal to zero, and zero otherwise.
ROA	Return on assets measured as income before extraordinary items at the beginning of the quarter.
$STDROA$	Standard deviation of ROA over the prior four fiscal quarters; $ASSETS$ is total assets in millions at fiscal quarter-end.
$NEGWORDS$	Percentage of negative words, as defined by the Negative Words dictionary of Loughran and McDonald (2011), in the entire conference call dialog less the percentage of negative words in the entire conference call dialog of the firm's prior-quarter earnings conference call.
$POSWORDS$	Percentage of positive words, as defined by the Positive Words dictionary of Loughran and McDonald (2011), in the entire conference call dialog less the percentage of positive words in the entire conference call dialog of the firm's prior-quarter earnings conference call.
$FREV$	Analyst one-quarter-ahead forecast revision, measured as difference between the median forecast for quarter $t+1$ earnings issued after and before the quarter t earnings announcement date, scaled by price 2 days before the earnings announcement. The median forecast before (after) the quarter t earnings announcement is measured as the last (first) forecast of all individual I/B/E/S analysts issuing forecasts during the 90-day period prior to (after) the quarter t announcement date.
$RECREV$	I/B/E/S summary consensus mean analyst recommendation revision, measured as the first I/B/E/S summary consensus mean after the conference call less the last I/B/E/S summary consensus mean before the conference call, where strong buy equals 5, buy equals 4, hold equals 3, sell equals 2, and strong sell equals 1.
$FDISP$	Standard deviation of analyst earnings per share forecasts.
$CAR(i,j)$	Daily abnormal returns cumulated over days i through j relative to the earnings conference call date, where expected returns are derived from the size and book-to-market portfolio to which the firm belongs.
UE_t	Unexpected earnings at period t measured as the difference between actual I/B/E/S earnings per share and I/B/E/S analyst summary consensus median earnings per share scaled by price per share 2 days before the conference call.
UE_{t+1} (UE_{t+2})	Unexpected earnings at period $t+1$ ($t+2$), whereas $UE_{t+1,t+2}$ is the aggregate unexpected earnings for $t+1$ and $t+2$. Measured as the difference between actual I/B/E/S earnings per share and I/B/E/S analyst summary consensus median earnings per share scaled by price per share 2 days before the conference call.
$LNMVE$	Natural logarithm of the market value of equity measured in millions at fiscal quarter-end.
MOM	Momentum measured as $CAR(-127, -2)$.
BM	Ratio of the book value of equity to the market value of equity at fiscal quarter end.
VOL	Stock return volatility measured as the standard deviation of daily stock returns over the period $(-127, -2)$ relative to the conference call date.

REFERENCES

Adler, Michael J., 2009, Detecting deceptive responses in sex offenders: A comparison of layered voice analysis (LVA) and the polygraph, Working paper, Counseling and Consultation Services, Inc. Limestone, TN.

Ambady, Nalini, and Robert Rosenthal, 1993, Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis, *Psychological Bulletin* 111, 256–274.

Andersen, Peter A., and Laura K. Guerrero, 1998, Principles of communication and emotion in social interaction, in P.A. Andersen, and L.K. Guerrero, eds. *Handbook of Communication and Emotion: Research, Theory, Applications, and Contexts* (Academic Press, San Diego, CA).

Antweiler, Werner, and Murray Z. Frank, 2004, Is all that talk just noise? The information content of internet stock message boards? *Journal of Finance* 59, 1259–1294.

Arnold, Magda B., 1960, *Emotion and Personality* (Columbia University Press, New York, NY).

Barsade, Sigal G., 2002, The ripple effect: Emotional contagion and its influence on group behavior, *Administrative Science Quarterly* 47, 644–675.

Bernard, Victor, and Jacob Thomas, 1989, Post earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* (Supplement), 1–36.

Bligh, Michelle C., and Gregory D. Hess, 2007, The power of leading subtly: Alan Greenspan, rhetorical leadership, and monetary policy, *The Leadership Quarterly* 18, 87–104.

Boersma, Paul, and David Weenink, 2010, PRAAT: Doing phonetics by computer (version 5.2.05) [Computer program], retrieved from http://www.praat.org/ on December 13, 2010.

Bradshaw, Mark, 2004, How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review* 79, 25–50.

Brown, Troy E., Stuart M. Senter, and Andrew H. Ryan, Jr., 2003, Ability of the Vericator to detect smugglers at a mock security checkpoint, Department of Defense Polygraph Institute report number DoDPI03-R-0002.

Bushee, Brian, Michael Jung, and Gregory Miller, 2011, Conference presentations and the disclosure milieu, *Journal of Accounting Research* 49, 1163–1192.

Cacioppo, John T., and Gary G. Bernston, 1994, Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates, *Psychological Bulletin* 115, 401–423.

Caffi, Claudia, and Richard W. Janney, 1994, Toward a pragmatics of emotive communication, *Journal of Pragmatics* 22, 325–373.

Collins, Daniel W., and S.P. Kothari, 1989, An analysis of the intertemporal and cross-sectional determinants of the earnings response coefficient, *Journal of Accounting and Economics* 11, 143–181.

Coval, Joshua D., and Tyler Shumway, 2001, Is sound just noise? *Journal of Finance* 56, 1887–1910.

Damphousse, Kelly R., Laura Pointon, Deidra Upchurch, and Rebecca K. Moore, 2007, Assessing the validity of voice stress analysis tools in a jail setting, Final Report to the United States Department of Justice: http://www.ncjrs.gov/pdffiles1/nij/grants/219031.pdf.

Davis, Angela K., Jeremy Piger, and Lisa M. Sedor, 2011, Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases, *Contemporary Accounting Research*, forthcoming.

Demers, Elizabeth A., and Clara Vega, 2010, Soft information in earnings announcements: News or noise? Working paper, INSEAD.

Diener, Ed., and Robert A. Emmons, 1985, The independence of positive and negative affect, *Journal of Personality and Social Psychology* 47, 1105–1117.

Ekman, Paul, and Wallace V. Friesen, 1974, Detecting deception from body or face, *Journal of Personality and Social Psychology* 29, 288–298.

Elkins, Aaron C., 2010, Evaluating the credibility assessment capability of vocal analysis software, *Proceedings of the Forty-Third Annual Hawaii International Conference on System Sciences*, January 5–8, 2010, Koloa, Kauai, Hawaii.

Elkins, Aaron C., and Judee K. Burgoon, 2010, Validating vocal analysis software to assess credibility in interpersonal interaction: A multilevel factor analytic approach, *Proceedings*

of the National Communication Association 96th Annual Convention, November 14–17, 2010, San Francisco, California.

Elliot, Andrew J., and Patricia G. Devine, 1994, On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort, *Journal of Personality and Social Psychology* 67, 382–394.

Engelberg, Joseph, 2008, Costly information processing: Evidence from earnings announcements, Working paper, Northwestern University.

Eriksson, Anders, and Francisco Lacerda, 2007, Charlatanry in forensic speech science: A problem to be taken seriously, *International Journal of Speech, Language and the Law* 14, 169–193.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Feldman, Ronen, Suresh Govindaraj, Joshua Livnat, and Benjamin Segal, 2010, Management's tone change, post earnings announcement drift and accruals, *Review of Accounting Studies* 15, 915–953.

Festinger, Leon, 1957, *A Theory of Cognitive Dissonance* (Stanford University Press, Stanford, CA).

Forgas, Joseph P., 2001, *The Handbook of Affect and Social Cognition* (Erbaum, Mahwah, NJ).

Frankel, Richard, Marilyn Johnson, and Douglas J. Skinner, 1999, An empirical examination of conference calls as a voluntary disclosure medium, *Journal of Accounting Research* 37, 133–150.

Freeman, Robert N., and Senyo Y. Tse, 1992, A nonlinear model of security price responses of unexpected earnings, *Journal of Accounting Research* 30, 185–209.

Frijda, Nico H., 1988, The laws of emotion, *American Psychologist* 43, 349–358.

Frijda, Nico H., 1993, Mood, emotion episodes, and emotions, in M. Lewis and J. Haviland, eds.: *Handbook of Emotions* (Guilford Press, New York, NY).

Gamer, Matthias, Hans-Georg Rill, Gerhard Vossel, and Heinz Werner Godert, 2006, Psychophysiological and vocal measures in the detection of guilty knowledge, *International Journal of Psychophysiology* 60, 76–87.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal, 2005, The economic implications of corporate financial reporting, *Journal of Accounting and Economics* 40, 3–73.

Han, Young, and Joseph Nunes, 2010, Please read the signal but don't mention it: How acknowledging identity signals leads to embarrassment, in *Advances in Consumer Research* 37 (Association for Consumer Research, Duluth, MN).

Harmon-Jones, Eddie, 2000, Cognitive dissonance and experienced negative affect: Evidence that dissonance increases experienced negative affect even in the absence of aversive consequences, *Personality and Social Psychology Bulletin* 26, 1490–1501.

Harnsberger, James D., Harry Hollien, Camilo A. Martin, and Kevin A. Hollien, 2009, Stress and deception in speech: Evaluating layered voice analysis, *Journal of Forensic Science* 54, 642–650.

Hatfield, Elaine, John T. Cacioppo, and Richard L. Rapson, 1994, *Emotional Contagion* (Cambridge University Press, New York, NY).

Hobson, Jessen, William J. Mayew, and Mohan Venkatachalam, 2011, Analyzing speech to detect financial misreporting, *Journal of Accounting Research* forthcoming.

Hollien, Harry, and James D. Harnsberger, 2006, Voice stress analyzer instrumentation evaluation, Final report (Contract: FA-4814–04-0011, CIFA, Gainesville, FL). Available at http://papers.ssrn.com/so13/papers.cfm?abstract˙id=1153450

Irvine, Paul, Marc Lipson, and Andy Puckett, 2007, Tipping, *Review of Financial Studies* 20, 741–768.

Johnson, Timothy C., 2004, Forecast dispersion and the cross-section of expected returns, *Journal of Finance* 59, 1957–1978.

Juslin, Patrick N., and Petri Laukka, 2003, Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin* 129, 770–814.

Juslin, Patrick N., and Klaus R. Scherer, 2005, Vocal expression of affect, in Jinni A. Harrigan, Robert Rosenthal, and Klaus R. Scherer, eds.: *The New Handbook of Methods in Nonverbal Behavior Research* (Oxford Press, Oxford, UK).

Kappas, Arvid, Ursula Hess, and Klaus R. Scherer, 1991, Voice and emotion, in B. Rimé and R.S. Feldman, eds. *Fundamentals of Nonverbal Behavior* (Cambridge University Press, Cambridge and New York).

Konopka, Lukasz, Jennifer Duffecy, and Kwan Hur, 2010, Technology assisted PTSD screening (TAPS) for Vietnam Veterans of America (VVA) study, pilot study abstract, The Chicago School of Professional Psychology, Northwestern University, and The University of Illinois at Chicago.

Lacerda, Francisco, 2009, LVA technology: The illusion of lie detection, in P. Branderud & H. Traunmuller, eds. *FONETIK 2009*, Department of Linguistics, Stockholm University.

Lazarus, Richard S., 1991, Progress on a cognitive-motivational-relational theory of emotion, *American Psychologist* 46, 819–834.

Lewis, Michael, 1993, The emergence of human emotions, in *Handbook of Emotions* (Guildford, NY).

Li, Feng, Michael Minnis, Venky Nagar, and Madhav V. Rajan, 2009, Formal and real authority in organizations: An empirical assessment, Working paper, University of Michigan, University of Chicago, and Stanford University.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries and 10-Ks, *Journal of Finance* 66, 35–65.

McNichols, Maureen F., and Patricia C. O'Brien, 1997, Self-selection and analyst coverage, *Journal of Accounting Research* 35, 167–199.

Mehrabian, Albert, 1971, *Silent Messages* (Wadsworth, Belmont, CA).

Mehrabian, Albert, and Morton Weiner, 1967, Decoding of inconsistent communications, *Journal of Personality and Social Psychology* 6, 109–114.

Neumann, Roland, and Fritz Strack, 2000, "Mood contagion": The automatic transfer of mood between persons, *Journal of Personality and Social Psychology* 79, 211–223.

O'Brien, Patricia C., Maureen F. McNichols, and Hsiou-wei Lin, 2005, Analyst impartiality and investment banking relationships, *Journal of Accounting Research* 43, 623–650.

Owren, Michael J., 2008, GSU PRAAT tools: Scripts for modifying and analyzing sounds using Praat acoustics software, *Behavior Research Methods* 40, 822–829.

Owren, Michael J., and Jo-Anne Bachorowski, 2007, Measuring emotion-related vocal acoustics, in *Handbook of Emotion Elicitation and Assessment* (Oxford University Press, New York, NY).

Palmatier, John J., 2005, Assessing credibility: ADVA technology, voice and voice stress analysis, in *Corporate Investigations* (Lawyers & Judges Publishing Company, Inc., Tuscon, AZ).

Piger, Jeremy, 2006, Is all that talk just noise? In *Monetary Trends*, Federal Reserve Bank of St. Louis.

Roseman, Ira J., 1984, Cognitive determinants of emotions: A structural theory, *Review of Personality and Social Psychology* 5, 11–36.

Scherer, Klaus R., 2003, Vocal communication of emotion: A review of research paradigms, *Speech Communication* 40, 227–256.

Scherer, Klaus R., Harvey London, and Jared J. Wolf, 1973, The voice of confidence: Paralinguistic cues and audience evaluation, *Journal of Research in Personality* 7, 31–44.

Schuller, Bjorn, 2010, On the acoustics of emotion in speech: Desperately seeking a standard, *Journal of Acoustical Society of America* 127, 1995.

Skinner, Douglas J., 2003, Should firms disclose everything to everybody? A discussion of "Open vs. closed conference calls: The determinants and effects of broadening access to disclosure," *Journal of Accounting and Economics* 34, 181–187.

Sommers, Mitchell S., 2006, Evaluating voice-based measures for detecting deception, *Journal of Credibility Assessment in Witness Psychology* 7, 99–107.

Sommers, Mitchell S., Troy E. Brown, Andrew H. Ryan, Jr., and Stuart Senter, 2007, Vericator: Assessing the validity and reliability of a voice-based method for detecting deception, Working paper, Washington University in St. Louis and Department of Defense Polygraph Institute, Fort Jackson, SC.

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

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance*, 63, 1437–1467.

Tomkins, Silvan S., 1962, *Affect Imagery Consciousness: The Positive Affects*, Volume 1 (Springer, New York, NY).

Ververdis, Dimitrios, and Constantine Kotropoulos, 2006, Emotional speech recognition: Resources, features and methods, *Speech Communication* 48, 1162–1181.

Walbott, Harold G., Pio Enrico Ricci-Bitti, and Eva Banninger-Huber, 1986, Non-verbal reactions to emotional experiences, in Klaus R. Scherer, Herald G. Wallbott, and Angela B. Summerfield, eds.: *Experiencing Emotion: A Cross-Cultural Study* (Cambridge University Press, Cambridge, UK).

Westphal, James D., and Michael Clement, 2008, Sociopolitical dynamics in relations between top managers and security analysts: Favor rendering, reciprocity and analyst stock recommendations, *Academy of Management Journal* 51, 873–897.

Wu, Chung-Hsien, Jui-Feng Yeh, and Ze-Jing Chuang, 2009, Emotion perception and recognition from speech, in *Affective Information Processing* (Springer Publishing, London, UK).

Yang, Bin, and Marco Lugger, 2010, Emotion recognition from speech signals using new harmony features, *Signal Processing* 90, 1415–1423.

Zuckerman, Miron, Mary D. Amidon, Shawn E. Bishop, and Scott D. Pomerantz, 1982, Face and tone of voice in the communication of deception, *Journal of Personality and Social Psychology* 43, 347–357.