2015-11-30

Contemporaneous verification of language: evidence from management earnings forecasts

Baginski, Stephen; Demers, Elizabeth; Wang, Chong; Yu, Julia

Contemporaneous verification of language: evidence from management earnings forecasts

Stephen Baginski
Terry College of Business, University of Georgia, Athens, GA 30602-6252, baginski@terry.uga.edu

Elizabeth Demers
Darden Graduate School of Business, University of Virginia, Charlottesville, VA 22906-6550, demerse@darden.virginia.edu

Chong Wang
Naval Postgraduate School, Monterey, CA 93943, cwang@nps.edu

Julia Yu
Nanyang Business School, Nanyang Technological University, Singapore 639798, juliayu@ntu.edu.sg

November 30, 2015

We thank workshop participants at University of Bocconi, University of Virginia (Darden), EDHEC, ESSEC, Exeter, IE Business School, the University of Notre Dame, the Stockholm School of Economics, the University of Verona, the 7th Interdisciplinary Workshop on Intangibles, Intellectual Capital & Extra-Financial Information in Warsaw, the 2012 Junior Faculty Accounting Research Symposium in Singapore, conference participants at the American Accounting Association (AAA) Annual Meetings in Denver, the European Accounting Association (EAA) Annual Congress in Rome, the University of Melbourne New Accounting Faculty Conference, and the 1st Accounting Conference at Catolica-Lisbon, Ben Ayers, Stephen V. Brown (AAA discussant), John Hassell, Karol Marek Klimczak (EAA discussant), Clive Lennox, Claudine Mangen, Shiheng Wang (University of Melbourne conference discussant), and Eric Yeung for helpful comments. We appreciate the programming assistance provided by Rajiv Jayaraman of Knolscape as well as Rashid Ansari. We thank Mary Boldrini, Pascale Gadroy, Cécile Maciulis, and Bingyun Wang for their research assistance and Craig Carroll for processing our text files using the Diction software. Demers is grateful for the research funding provided by the INSEAD Alumni Fund. Earlier versions of this study were circulated under the title "Understanding the Role of Language in Management Forecast Press Releases."

Abstract

Research documents that linguistic tone is incrementally informative about stock returns. What remains a puzzle is the mechanism by which investors can assess its credibility. We examine whether contemporaneous information in management earnings forecasts serves as a timely alternative to ex post verification. We document that ex post verifiable quantitative news in unbundled forecasts, and characteristics of the linguistic tone itself, affect investors' pricing of tone. Consistent with higher quality signals enhancing the credibility of contemporaneous lower quality signals, we find that quantitative news verifies the associated linguistic tone; when the two signals have the same sign, the price effect of tone is stronger. Furthermore, the pricing attenuation of tone is increasing in the imprecision of the quantitative forecast, suggesting that lower forecast quality compromises the quantitative signal's credibility enhancement. Managerial incentives to inflate tone lead to the verification effect being greater for optimistic language, while management's use of hyperbole results in attenuation of the tone's pricing.

JEL Classifications: G14; D82; M41

Keywords: Management forecasts; linguistic tone; cheap talk; management incentives

1 Introduction

A growing body of literature seeks to understand the properties and price relevance of language (often referred to "linguistic sentiment" or "linguistic tone") in corporate disclosures, incremental to simultaneously released quantitative news. The null hypothesis in these studies, that language is *not* incrementally priced, is hardly a straw man, given that language shares several of the key characteristics of cheap talk; notably, it is costless to convey and relatively difficult to verify, even ex post. However, in spite of its propensity to be cheap talk, studies continue to document that language is priced in several situations: earnings announcements (Davis et al. 2012; Demers and Vega 2014), restatement announcements (Mangen and Durnev 2010), IPO prospectuses (Balakrishnan and Bartov 2010), conference calls (Chen et al. 2014), and MD&A and other elements of the 10K reports (Davis and Tama-Sweet 2012; Feldman et al. 2010; Li 2010).

Although prior research documents language's incremental pricing, it does not identify a mechanism that serves as a timely alternative to ex post verification by which investors can assess the likely credibility of language. Our objective is to begin to fill this gap in the literature by investigating whether information in the disclosure itself provides a contemporaneous verification of the credibility of language and thus strengthens its market pricing. To accomplish this, we examine the pricing of linguistic tone in management forecast disclosures that are not bundled with earnings releases. These disclosures possess certain properties that provide a relatively well-controlled environment in which to test our contemporaneous verification conjecture. In unbundled management forecast disclosures, two primary signals are present: the quantitative (i.e., hard) management forecast news and linguistic tone.[1] One of the two signals, the quantitative management forecast, is verified by the subsequent earnings realization. Stocken (2000) examines the credibility of a manager's disclosure of information in a repeated cheap talk

[1] Management forecasts bundled with earnings releases (e.g., Rogers and VanBuskirk 2012) may confound the analysis with a third signal, the earnings realization. Earnings releases also contain a substantial amount of new disclosures that can affect stock prices and that make it difficult to link linguistic tone to a given hard signal, adding noise to the analysis. For these reasons, we discard management forecasts bundled with earnings releases to perform our main tests but include them later in a supplemental test.

game and concludes that the manager almost always truthfully reveals private information under certain sufficiency conditions, including an accounting report that can be used to assess credibility ex post. Thus, because it is verifiable, the quantitative management forecast is expected to be a higher quality signal than the accompanying linguistic tone. However, because managers choose the precision of their quantitative forecasts, the quality of quantitative forecasts varies in the cross-section.

We investigate three factors that affect investors' pricing of linguistic tone: 1) the correspondence between contemporaneously observable, ex post verifiable quantitative management forecast news and the ex post less verifiable linguistic tone; 2) the precision of the quantitative management forecast; and 3) a characteristic of the linguistic tone, the use of hyperbole. First, we predict that quantitative management forecasts serve as contemporaneous verifiers of accompanying linguistic tone; when the two signals are consistent (i.e., the signs of the quantitative management forecast and linguistic tone agree), linguistic tone is priced more. Past empirical results in other settings, as well as economic models, suggest that, in the presence of two signals, the higher quality signal can add credence to the potentially lower quality signal. For example, Jennings (1987) documents that contemporaneous financial analyst revisions serve to verify the veracity of management forecasts. Baginski and Hassell (1990) document that analyst forecast revisions are greater when the market reaction to a management forecast agrees with the management forecast news. Hutton et al. (2003) document that a potentially less credible good news management earnings forecast is perceived as more credible when it is issued with more verifiable information. Dambra et al. (2013) document that less verifiable non-GAAP forecasts are viewed by investors as more credible if they are accompanied by a verifiable management earnings forecast. The notion that a verifiable, co-existing public signal affects the precision and information of a second signal based on a sender's private information is also consistent with the cheap talk model proposed by Fischer and Stocken (2010). Second, we predict that the verification effect is stronger when a good news quantitative forecast confirms positive linguistic tone, consistent with the notion that good news is inherently less credible than bad news in the voluntary disclosure setting (when the signal is less ex post verifiable), due to management incentives to increase stock prices. Third, we predict that the ability of the quantitative management forecast to contemporaneously corroborate linguistic tone is compromised by lower quantitative forecast quality (e.g., imprecise forecast forms and rounding as identified by Baginski et al. 1993; Bamber and Cheon 1998; Bamber et al. 2010), and thus linguistic tone pricing is attenuated when management forecasts are imprecise. Fourth, we predict that linguistic tone pricing is attenuated by a characteristic of the tone itself, management's use of extreme language (i.e., hyperbole) (Admati and Pfliederer 2004).

Using a Factiva search, we identify a sample of 1,754 management earnings forecasts from 1997 through 2006 from which we derive a measure of linguistic tone. After empirically confirming two results from prior language research in our unique forecast setting—that is, that linguistic tone is incrementally priced and that tone is generally consistent with the accompanying hard news—we find results that support our four predictions. Our findings are not confounded by the presence of external attributions (Baginski et al. 2004), other disclosure content including special items that accompany forecasts, or limiting our sample to unbundled forecasts. Furthermore, we find no evidence of short-run reversal of the verification effect in post-announcement drift tests.

Our study extends the linguistic tone literature in several ways. First, we establish that quantitative management forecasts serve as contemporaneous verifiers of linguistic tone (more so for optimistic tone, consistent with management incentives to inflate prices), and we thereby provide evidence of a mechanism by which investors contemporaneously assess the veracity of tone information that is otherwise difficult to verify, even ex post. Second, we document that hyperbole serves as a contemporaneous modifier of linguistic tone, consistent with the predictions of theoretical cheap talk models. Although not the primary motivation for our study, we also confirm that the economically meaningful pricing effect associated with linguistic tone that accompanies other types of quantitative disclosures (e.g., earnings realizations, quantitative explanations of performance in the MD&A, restatements, and conference calls) extends to the setting of voluntarily provided, unbundled management forecast announcements. In addition to being confirmatory, our evidence offers a unique contribution to this prior set of studies. On the one hand, one might expect that the imprecise language signal would have greater information content in the management forecast situation vis-à-vis the earnings release setting because it does not have to compete with the substantial amount of information in the earnings release setting, much of which is subject to auditor scrutiny. On the other hand, the unbundled management forecast is primarily an incentives-driven disclosure event. Given that many of these incentives may lead to lower quality disclosures of both expected earnings and language, there is substantial tension about whether language is credible in the management forecast setting and whether the management earnings forecast (itself incentives-driven) can verify language. Furthermore, because the primary underlying economic reason to expect linguistic tone to be informative for prices is its predictive power for earnings, it is most surprising to find that language remains statistically and economically significant when it accompanies management-issued earnings *forecasts*.

Our findings also extend the understanding of the pricing consequences of management forecasts. Prior management forecast research documents the effects of supplemental disclosures on the pricing of quantitative forecast news (Hutton et al. 2003; Baginski et al. 2004). We complement this finding by documenting the reverse effect of quantitative management forecast news on the pricing of supplemental linguistic tone. We also show that hard management forecast precision, as measured by forecast form and rounding, has implications beyond its effects on the pricing of the hard forecast news itself; forecast precision also is associated with the market pricing of accompanying language.

In the sections that follow, we develop our three hypotheses, describe data and the research design, and then present results and supplemental empirical tests.

2 Hypothesis development

2.1 The agreement hypothesis

The price-relevance of language is plausible if language is a sufficiently reliable signal of future earnings or dividends. However, language is relatively costless to provide, difficult to verify, and likely linked to future earnings and dividends in a fairly noisy way. The difficulty of ex post verification, in particular, calls the informativeness of cheap talk into question (Crawford and Sobel 1982; Benabou and Laroque 1992; Dye and Sridhar 2004). However, prior empirical work shows that, *on average*, linguistic tone is incrementally informative for security prices when released with an earnings announcement (Davis et al. 2012; Demers and Vega 2014) and in other communications (e.g., IPO prospectuses, restatement announcements, and conference calls). These studies also document an association between linguistic tone and future earnings and between tone and the uncertainty of future earnings, thus establishing plausible links to valuation fundamentals as the reason for the information content of language. The information content of linguistic tone is also suggested by a general aversion to lying (Gneezy 2005; Hurkins and Kartik 2009). Furthermore, models in behavioral economics (e.g., Mullainathan et al. 2008) and experimental results (e.g., Bertrand et al. 2010) suggest that uninformative material can affect choice, while archival evidence also supports the conclusion that apparently uninformative disclosures can matter (Michels 2012).

The language that we investigate accompanies management earnings forecasts. A manager provides two potentially price-relevant signals, one a quantitative forecast of earnings, the other a set of words from which linguistic tone can be ascertained. On average, we expect the quantitative management forecasts to be credible. Ajinkya and Gift (1984) provide evidence that managers use forecasts to adjust market expectations by revealing their private information, and King et al. (1990) provide the economic underpinnings of the expectations-adjustment hypothesis. Stocken (2000) examines the credibility of a manager's disclosure of information in a repeated cheap talk game and concludes that the manager almost always truthfully reveals private information under certain sufficiency conditions, including an accounting report that can be used to assess credibility ex post (i.e., as is the case for the quantitative management forecast) and the ability to make a longer-run assessment of management credibility.

In summary, the verifiability of quantitative management forecasts enhances their credibility, and thus they can serve as contemporaneous verifiers of the accompanying linguistic tone. If the intent of a management forecast press release is to adjust market expectations through the credible revelation of private information, then it is less likely that the manager would issue conflicting linguistic tone and quantitative forecast news. Indeed, the general agreement of linguistic tone and hard earnings news in the earnings press release setting is verified empirically by Demers and Vega (2014), and we document that it is commonly the case in our forecast setting as well.

We expect the price relevance of linguistic tone to be strengthened by the agreement of the quantitative management forecast with the linguistic tone. Our view is consistent with that of Jennings (1987), who documents that contemporaneous financial analyst revisions verify the veracity of management forecasts; with that of Baginski and Hassell (1990), who document that analyst forecast revisions are greater when the market reaction agrees with management forecast news; with that of Hutton et al. (2003), who argue that less believable news is more credible when issued with more verifiable information; and with that of Dambra et al. (2013), who document that less verifiable non-GAAP forecasts are viewed by investors as more credible when accompanied by a verifiable management earnings forecast. The notion that a credible, co-existing public signal affects the precision and informativeness of a second signal based upon the sender's private information is also consistent with the cheap talk model proposed by Fischer and Stocken (2010). They find that, when information costs are moderate, more precise public information leads an incentives-driven information sender to gather more precise private information to increase the decision-maker's use of the disclosure and facilitate credible communication. Stated in the alternative form, our first hypothesis (the "agreement" hypothesis) is therefore as follows:

H1a The price response to linguistic tone accompanying a quantitative management forecast is increasing in the agreement of the linguistic tone and the management forecast news as measured by whether their signs agree (i.e., linguistic tone is optimistic and management forecast news is good or it is pessimistic and management forecast news is bad).

The fact that management forecasts are also incentives-driven adds tension to this hypothesis. That is, the hypothesis is predicated on the credibility of management earnings forecasts. We predict that the more precise signal (the management earnings forecast) verifies the less precise one (language). We have no basis for hypothesizing the reverse.

2.2 The management incentives hypothesis

Management incentives play a role in the strength of our H1a prediction. We expect the effect of sign agreement predicted in H1a to be greater when the quantitative management forecast news sign is positive and confirms an optimistic tone (relative to when negative forecast news confirms a negative tone). This additional conjecture is supported by cheap talk theories, which suggest that language inflation (i.e., overly optimistic language) occurs and that this phenomenon exists regardless of the costs of lying (Kartik 2009). Although a rational receiver of information is not deceived and thus discounts the language, the sender nevertheless does not remove the inflation because the sender fully expects the receiver to apply discounting. Blanes i Vidal (2003) modifies the model proposed by Morgan and Stocken (2003) to establish an equilibrium in which investors react more to bad news than to good news in the presence of cheap talk. The receiver's presumption that language is inflated is driven by the relatively few management incentives to release bad news and numerous incentives to increase share prices.[2] If management incentives lead to language inflation and inflation attenuates price response to net

[2] This asymmetry does not necessarily arise from managers' opportunism. Many of the managers' personal incentives are aligned with the firm's incentives via share ownership, compensation based on firm performance, and the shared benefits of a good reputation. Even in the relatively few cases where the manager's motive for voluntary disclosure is opportunistic, only a handful of situations suggest purely self-serving incentives to downward bias disclosure news. For example, bad news might be strategically timed before the date of determining the share price for an option grant (Aboody and Kasznik 2000), before a management takeover attempt (Hafzalla 2009), and before insider buying. This is not to say that these events do not occur or are not economically meaningful to the parties involved, just that one would expect that, in the broad cross-section, they are not sufficiently pervasive to suggest an expectation of language deflation.

positive tone, we predict that the contemporaneous verification effect is potentially strongest for net positive relative to net negative tone:

H1b The effect of sign agreement is greater for optimistic (relative to pessimistic) linguistic tone.

This effect is driven by the expected optimistic bias of the *linguistic tone* not of the management forecast. The theory underlying the H1b prediction treats language, being largely unverifiable, as cheap talk. Theory predicts a discounting of good-news cheap talk relative to bad-news cheap talk when incentives for good-news disclosure exist. Significant tension is introduced if good-news *management forecasts* (i.e., the verifying signals) are less credible than bad-news management forecasts. However, we expect good-news management forecasts to be equally credible verifiers of language because management forecasts possess a characteristic, ex post verifiability, that is not a characteristic of cheap talk (Stocken 2000).[3]

2.3 The hard news precision hypothesis

Although more *ex post* verifiable than language, the ability of a quantitative management forecast to serve as a contemporaneous verifier varies in the cross-section because management forecasts are often imprecise and thus of lower quality. Management forecasts are often expressed in ranges, in minimums or maximums, or rounded to the nearest nickel (Baginski and Hassell 1997; Bamber and Cheon 1998; Bamber et al. 2010). If management forecast imprecision hinders the forecast's verification role, it decreases the price impact of accompanying linguistic tone (the hard-news precision hypothesis):

H2 The price response to linguistic tone accompanying a quantitative management forecast is decreasing in the imprecision of the management forecast as measured by the width of the forecast range and whether the forecast is rounded.

[3] Consistent with the role of ex post verifiability as a disciplining mechanism for management earnings forecasts, Rogers and Stocken (2005) and Kothari et al. (2009) find that good-news management forecasts are not less accurate than bad-news management forecasts. Merkeley et al. (2013) also suggest that good-news management forecasts are not less credible than bad-news management forecasts.

2.4 The hyperbole hypothesis

Finally, we consider a contemporaneously available characteristic of the language itself. Hyperbole is the use of language in the extreme to provide emphasis, even though the expressions are not to be taken literally. In addition to the frequent use of hyperbole in creative writing, hyperbole manifests in evaluative contexts such as online reviews of movies, products, and physicians. For example, Anderson (1998) argues that the consumer's marginal utility from word of mouth rating of a product increases with the amount of either satisfaction or dissatisfaction, yielding the expectation and empirical finding of a U-shaped customer satisfaction distribution. Using a cheap talk model, Admati and Pfliederer (2004) demonstrate that, if the underlying distribution is uniform, sender overconfidence always leads to a decrease in the amount of information transmitted in equilibrium because of the sender's tendency to exaggerate. More extreme messages characterize the equilibrium.

To the extent that hyperbole bias exists, either due to its marginal utility or as a consequence of overconfidence, we expect extreme net positive and net negative tone to be discounted by investors (the hyperbole hypothesis):

H3 The price response to linguistic tone accompanying a quantitative management forecast is decreasing in the absolute magnitude of linguistic tone.

In summary, our hypotheses are based solely on observable contemporaneous characteristics in the management earnings forecast disclosure: the sign of the quantitative forecast news, the width of the quantitative forecast range, whether the forecast is rounded, and the sign and absolute magnitude of the linguistic tone measure. In tests described later, we control for a host of variables capturing other potential cross-sectional determinants of the price reaction to linguistic tone.

3 Sample and language data

3.1 Main sample determination

We use the Factiva database to individually identify and download candidate management earnings forecasts. We follow Baginski et al. (2004) by using business newswires Dow Jones Business News ("DJBN") and Press Release Newswire ("PRN") to search for the following word strings—"expects earnings," "expects net," "expects income," "expects losses," "expects profits," and "expects results"—in addition to three parallel lists where "expects" is replaced alternatively by "forecasts," "predicts," and "sees". As reported in Table 1, this search yielded 6,180 candidate earnings forecasts (3,577 for DJBN and 2,603 for PRN) for the period of 1997 through 2006. We next created individual .txt files for each candidate management forecast article and extracted firm identifiers for the companies underlying each respective Factiva article to match the candidate observations into the CRSP, Compustat, and First Call databases.[4] To be included in our initial sample, each candidate management forecast from Factiva had to match up with a management forecast from the First Call Company Issued Guidance database within a three-day window surrounding the Factiva date.[5] We delete 2,136 observations with fewer than 100

[4] Where available, the ticker symbol associated with the firm covered in the newswire article was extracted electronically from the text of the Factiva download using the Factiva ticker symbol tag. In many instances, particularly for the PR Newswire articles later in the sample period, there were no tagged ticker symbols included in the Factiva output. Accordingly, we developed an algorithm that used the unique Factiva codes to generate smart guesses at the firm's ticker symbol. Factiva has assigned unique ticker-symbol-like codes to each firm in its database. These codes are often similar, if not identical, to the firm's ticker symbol. Accordingly, we first tried to match the Factiva code as closely as possible to a ticker symbol in the CRSP master list using Stata. We then verified the accuracy of these guesses by using a second automated algorithm to compare a string from the Factiva tagged company name to a string from the company name in the CRSP master list. All of these automated "candidate matches" were then manually verified. Most of the incorrect matches related to companies whose names contained initials or that were characterized as having a generic first name in their corporate name (e.g., "General" or "National"). In over 800 cases, a Factiva code, ticker symbol, or company name was not extractable using standard programming techniques. For these unmatched Factiva observations, as well as for those that were found to have been erroneously matched due to initialed or generic corporate names, we manually matched the Factiva article's company details into the CRSP master list. A review of the observations we failed to match suggests that many of them relate to foreign companies, venture capital organizations, pre-IPO entities, or other firms that are appropriately excluded from our sample, or that the Factiva articles were generated by market analysts discussing a sector or firm's prospects rather than being management forecasts.

[5] The requirement that our press release observations match with First Call earnings forecasts results in the loss of some press releases. Chuk et al. (2013) document that excluded observations are likely to be made by firms that a) have less analyst forecast coverage, b) have poor prior performance, c) do not issue an accompanying earnings release, and d) do not provide an EPS forecast (e.g., provide a revenue or cash flow forecast instead). Because we require analyst forecasts for our observations, have other data requirements that skew our sample toward larger firms, focus on earnings forecasts, and sample primarily after 1997 (which is suggested by Chuk et al. (2013) as a means of mitigating bias), we do not believe that our sample would be meaningfully different had First Call matches not been required. However, the data required by our research design does result in a larger firm sample bias, consistent with that characterizing many large sample capital markets studies.

words in the forecast announcement or for which required data is not available on First Call Management Issued Guidance database (forecast characteristics and analyst forecasts), Center for Research in Security Prices (CRSP) (return and price data), and Compustat (accounting data).

To increase the internal validity of our design, we also discarded 2,035 forecasts bundled with earnings releases. Following the literature (e.g., Anilowski et al. 2007; Rogers and Van Buskirk 2012), we define *bundled* management forecasts as those falling within two days of an earnings announcement date. Rogers and Van Buskirk (2012) argue that the sign of the unexpected hard management forecast news is potentially measured incorrectly when accompanied by an earnings release (i.e., a bundled forecast) because the analyst expectation used to compute the unexpected management forecast news is out of date; that is, it has not been updated in light of the implications of the accompanying earnings release for analysts' expectations of future earnings. Earnings announcements also contain a host of additional price-relevant disclosures for which control is difficult. Furthermore, it is not clear to which of these additional disclosures the linguistic tone in the announcement pertains. Finally, we read the remaining management forecast disclosures and discarded 245 because they referred to the management forecasts of more than one firm or were ex post reports about the price reactions to the management forecasts. Our final sample consists of 1,764 management earnings forecasts issued by 750 distinct firms over our 10-year sample period. Typical of prior management forecast studies, our sample is highly diversified across firms and time. It contains annual and quarterly forecasts; so, if firms issued forecasts in each quarter and year, we would observe 37,500 forecasts (750 firms x 10 years x (4 quarters + 1 annual forecast per year). Less than 5 percent of those potential periods contain an *unbundled* management earnings forecast, with the large majority of firms in our sample issuing only one or two *unbundled* management forecasts during the sample period (i.e., 336 issued one forecast, and 223 issued two forecasts).

3.2 Measuring the language construct

We obtain our measure of linguistic tone using textual analysis of management earnings forecast disclosures.[6] Prior evidence (Loughran and McDonald 2011; Demers and Vega 2014) suggests that generic linguistic algorithms such as Diction or General Inquirer may yield noisy measures of "positive" and "negative" linguistic tone in the context of financially oriented text passages. Accordingly, we use the Loughran and McDonald (2011) (L&M) finance-oriented dictionaries (i.e., word lists) for capturing "positivity" and "negativity" (measures that L&M refer to as Fin-Pos and Fin-Neg, respectively).[7] Following the financial linguistics literature (e.g., Davis et al. 2012; Demers and Vega 2014), our principal measure of tone is defined as the difference between positive (optimistic) and negative (pessimistic) words in the disclosure deflated by number of total words in the disclosure, which we label as *NetPositivity*.

Similar to most prior studies, our *NetPositivity* variable does not attempt to explicitly measure the "unexpected" portion of tone. Although some studies have examined the time-series properties of tone in earnings releases, leading them to adopt the *change* in net optimism as a proxy for the "unexpected" or "news" component of linguistic tone, we have opted not to do so for several reasons. First, as Loughran and McDonald (2011) point out, measuring "unexpected" tone in this way imposes a considerable amount of structure on the linguistic parameters; specifically, it presumes a considerable amount of processing capability on the part of investors in the cross-section. While this is conceivable in the context of

[6] We follow the cleansing procedure proposed by Demers and Vega (2014). Corporate press releases on the newswire services often include several paragraphs at the end of the announcement that are not part of the body of the announcement that matters to our study. Specifically, the releases typically include a company-standard paragraph that describes the firm, often using very flattering language. In addition, most of the articles tend to include some form of safe-harbor disclaimer related to the forward-looking information included in the press release. These latter paragraphs vary somewhat across firms but are generally boilerplate and are presumed to be drafted by the company's lawyers. Finally, the press releases typically end with company contact information such as a listing of the corporate website address, their investor relations contact names and numbers, information related to upcoming conference calls, or a combination of these. Since all of these paragraphs contain textual and numerical data that may include language relating to optimism, pessimism, and certainty that does not form part of the content portion of management's press release per se, based upon manual review and the identification of keyword strings, we developed algorithms to cleanse the .txt files of such non-announcement content. These cleansed .txt files are used as the basis for the linguistic characteristics extracted from the firm's earnings forecast press releases.

[7] Rogers et al. (2011) argue that more general language measures such as Diction might be more relevant in the context of communications with nonfinancial audiences such as attorneys, judges, and others. In untabulated analyses described later, we find that, in our sample, the Diction-based linguistic tone measure is not as price-relevant as the L&M measure and is not price relevant incremental to the L&M measure.

regularly recurring, mandated, quarterly earnings announcements, it is much less likely to hold in the context of sporadically issued, nonstandardized management forecasts. Second, Demers and Vega (2014) take this more refined approach (in the earnings release context) and report that their basic findings are unaffected by the use of net optimism rather than the change in net optimism. Finally, a sufficient time series of management forecasts is not available for the majority of our sample firms, precluding the modeling of the time-series behavior of tone in the context of management forecasts. The consequences of not specifying an expectations model for our tone variable would generally be to introduce noise into the measure, thereby reducing the power of our cross-sectional tests.[8] As we discuss below, however, the results in relation to the *NetPositivity* variable are generally quite strong and significant.[9]

4 Empirical results

4.1 Descriptive statistics

Table 2 presents descriptive statistics for the forecasts, linguistic tone, and other regression variables.[10] In Panel A, consistent with prior research, the unbundled management forecasts in our sample are, on average, not good news, as evidenced by the negative means and medians for abnormal announcement returns (*AR)* and the near zero mean and median management forecast surprises (*FSURP)*.[11] Mean *NetPositivity* is similarly slightly negative, and the median is zero. Panel B shows that the majority of

---

[8] Related to this point, introduction of an expectations model adds additional noise to the analysis if linguistic tone is a change rather than a levels variable. It is not clear that *NetPositivity* is a levels variable. For example, words like "better" and "improve" are positive words that indicate a change in prospects rather than a level.

[9] The text that we use is from the financial press or newswire coverage of a firm's earnings guidance press release. The text passages largely reiterate the management's filing. We recognize that management support personnel (e.g., investor relations and legal professionals) are likely to influence the filing. Therefore one can view the disclosure as generated by a management team. Regardless of whether the language is individual-generated or team-generated, it is still substantially less ex post verifiable. Furthermore, reporter or editor choices to paraphrase rather than directly quote might add noise and make it more difficult to reject the null hypotheses in H1b (the discretion-related hypothesis) in the direction we expect.

[10] The distributions presented in Table 1 are after winsorizing every continuous, naturally unbounded variable at both the 1st and 99th percentiles.

[11] *AR* is the size- and book-to-market-adjusted three-day announcement return centered on the announcement date t=0. *FSURP* is the management forecast surprise measured as the management forecast minus the preceding median consensus analyst forecast, all scaled by pre-announcement (t=−2) security price. To compute the management forecast for nonpoint forecasts, we use the midpoint of ranges and the disclosed upper or lower bound for maximum and minimum forecasts, respectively, all consistent with the prior management forecast literature.

forecasts convey quantitative news that agrees in terms of sign with the sign of *NetPositivity* (66.4%), are not rounded (74.6%), occur after Regulation FD (61.3%), and occur before the Sarbanes-Oxley Act (65.2%). Panel C shows that the forecasts are primarily range forecasts (70.2%). Panel D shows a high dispersion of forecasts across years, with an initial increase in forecast incidence over time followed by a decrease starting around the time of Regulation Fair Disclosure (Reg FD) in late 2000 and Sarbanes-Oxley (SOX) in 2002.[12]

4.2 Confirmation of prior findings in the management forecast setting

The purpose of Table 3 is to confirm two findings in the prior literature for our sample of unbundled management forecasts: 1) linguistic tone is price relevant incremental to hard news (in our case, the quantitative management forecast) surprise, and 2) the signs of the hard news and the simultaneously released linguistic tone tend to agree. To confirm the pricing relation, we estimate the following pooled cross-sectional, time-series ordinary least squares model:

$$AR_{jt} = \beta_0 + \beta_1\, FSURP_{jt} + \beta_2\, NetPositivity_{jt} + \beta_3\, OtherFSURP_{jt} + \varepsilon_{jt}, \tag{1}$$

where j is a firm subscript, t represents a management forecast at date t, *OtherFSURP* is the value of *FSURP* for any other management forecast in the press release and set equal to zero if another forecast is not present, and all other variables are as previously defined. $\beta_2 > 0$ provides confirmation of the findings of prior literature in our management earnings forecast sample.[13]

Two key research design issues are relevant to the estimation of equation (1). The first issue is the degree of observation independence. Our sample is highly diversified in event time and by firm. That is, typical of management forecast samples, there are very few firms releasing on any given date, suggesting

[12] Kross, Ro, and Suk (2011) report means for the First Call population for 1996–2008. Based on their analysis, their (our) means on two key descriptors of the sample—log of firm size and log of one plus analyst following—are 6.70 (7.22) and 2.40 (2.00), respectively. Thus our final sample is relatively similar to the First Call population of firms.

[13] Ng, Tuna, and Verdi (2011) document a delayed response to unexpected quantitative management forecast news. Although market efficiency is a maintained assumption in our analysis, given the post-announcement drift literature and the Ng et al. (2011) finding, we also discuss post-announcement results in supplemental tests described later.

that mean cross-sectional correlation is very low, and, although a given firm might enter the sample more than once, the *AR* measurements are not consecutive. As a result, there is no reason to believe that serial correlation in the dependent variable (and thus in the residuals) exists. For example, Bernard (1987) demonstrates very low serial correlation in consecutive *AR* (again, ours are not consecutive and thus the expected correlation would be even lower) and little effect on standard errors of ignoring serial correlation. However, there may be two management forecasts in a single announcement. When this occurs, *AR* is measured at the same date for both observations, and thus the two observations are not independent. Accordingly, we cluster these observations and control for year fixed effects so that all reported tests are based upon standard errors computed from independent observation clusters (see Petersen 2009).[14] Our clustering approach yields 1,157 clusters (approximately 1.5 observations per cluster on average).

The second issue in estimating equation (1) is control for other information. Inclusion of *OtherFSURP* controls for the other earnings information in the press release. However, other non-earnings information can exist in a management forecast announcement. We consider this issue in supplemental tests described later.

Table 3 presents the results from estimating equation (1). We present the results for regressions on hard forecast news (*FSURP*) and linguistic tone (*NetPositivity*) separately and then together to confirm the incremental information content of linguistic tone. The results support the incremental pricing of

---

[14] As pointed out by Peterson (2009), the appropriate clustering procedure (or fixed effect modeling) in panel data depends on whether one has reason to believe that dependence exists in the residuals in the cross-section (date) and time-series (firm). Gow et al. (2010) make this point as well. In describing Peterson's findings in two finance applications, asset pricing and capital structure, they state: "He concludes that in these settings 'clustering standard errors by both firm and time appears unnecessary' (Peterson 2009, 473). In contrast, we find that, in a variety of accounting applications, two-way cluster-robust standard errors are required for valid inferences. The difference between our findings and those of Peterson (2009) is due to the fact that accounting variables (e.g., earnings and credit ratings) exhibit greater dependence both over time and in cross-section than finance variables (e.g., returns)." We use returns in our tests. An alternative approach would be to discard one of the forecasts from the sample. But, because each forecast has its own properties of interest (e.g., sign, form, rounding, horizon), we do not follow this approach in our main tests. In an untabulated analysis, we discard one of the forecasts, repeat our tests, and obtain consistent results. Finally, although we believe that two-way clustering is not necessary, to ensure the robustness of our analysis, in untabulated tests, we use alternative clustering approaches including two-way clustering (described later).

linguistic tone. The coefficients on *FSURP* and *NetPositivity* are both significantly positive in their individual regressions and when considered jointly.[15]

The incremental economic effect of a change in *NetPositivity* on security prices is substantial. In the multiple regression, a one standard deviation increase in *FSURP* is associated with an increase in three-day size and book-to-market adjusted returns of 167 basis points, and a one standard deviation increase in *NetPositivity* is associated with an increase in adjusted returns of 336 basis points (results not tabulated). The relative economic importance of linguistic tone in our sample of management forecast press releases is initial evidence of a stronger role for language in purely voluntary disclosure settings. Demers and Vega (2011) report an increase in returns for a one standard deviation of linguistic tone of 72 basis points in the earnings release setting.[16]

Table 4, Panel A, presents a chi-square test of whether the signs of linguistic tone and the hard management forecast news tend to agree. The hard forecast news sign is good (bad) news when *FSURP* ≥ 0 (< 0) and the linguistic tone proxy is good (bad) when *NetPositivity* ≥ 0 (< 0). Examination of the main diagonal reveals that 66.4% of the forecasts in our sample contain linguistic tone that is directionally consistent with the hard earnings forecast surprise. In other words, good news forecasts tend to be accompanied by net optimistic language, while bad news forecasts are predominantly accompanied by net pessimistic language. Chi-square tests reject the null of no association between the signs of forecast news

---

[15] The coefficient estimate on *OtherFSURP* is not significant in the expected positive direction. *OtherFSURP* is insignificant because it is largely redundant when present and not always present. When considered in isolation in a simple regression, *OtherFSURP* has a t-statistic of 3.0. Also, our sample includes forecast releases occurring close to the period end, which some studies refer to as "warnings." Our conclusions are not affected by discarding management forecasts issued less than 14 days before the earnings release date. As an alternative specification check, we modify equation (1) to include intercept and slope shifts for these warnings. The insignificant coefficients obtained on these variables indicate that the price reactions to quantitative forecast and linguistic tone do not differ for warnings. We also replicate our analysis using the alternative Diction linguistic tone proxy. Consistent with greater noise in this proxy, its t-statistic is much smaller in a simple regression. When both proxies enter the multiple regression, the Diction proxy is subsumed by the L&M proxy. That is, the L&M proxy is strongly significant and the Diction proxy is insignificant. This relative strength of the L&M proxy is consistent with prior research findings.

[16] If we rerun our analysis for management forecasts bundled with earnings releases to come closer to the Demers and Vega sample, we find an increase in returns for a one standard deviation of linguistic tone of 94 basis points for the L&M proxy and 48 basis points for the Diction proxy, an average of 71 basis points. The 72 basis point effect in Demers and Vega (2011) is obtained using a linguistic tone proxy that combines the L&M proxy, the Diction proxy, and a third proxy.

and linguistic tone. Although these findings are consistent with the results of prior linguistic studies that have documented a uniformly positive correlation between linguistic tone and unexpected earnings in earnings press releases, they contrast sharply with a general finding in prior research of strategic bundling of good news with bad news, presumably to reduce the price consequences of bad news (e.g.,Waymire 1984; Rogers and Van Buskirk 2012). We note, however, that the sign association is not perfect; signs disagree in 33.6% of the cases, and the off-diagonals are roughly equally divided between positive and negative linguistic tone. Table 4, Panel B, documents significantly positive Pearson product-moment and Spearman rank-order correlations between the continuous *FSURP* and *NetPositivity* variables.

4.3 Contemporaneous verification of linguistic tone

To examine our main hypotheses related to the contemporaneous verification of linguistic tone, we estimate the following model:

$$AR_{jt} = \gamma_0$$

Baseline: $+\ \gamma_1\, NetPositivity_{jt}$

Agreement (**H1a**)/Management incentives (**H1b)**:

$$+\ \gamma_2\, SignsAgree_{jt} + \gamma_3\, SignsAgree{+/+}_{jt}{*}NetPositivity_{jt} + \gamma_4\, SignsAgree{-/-}_{jt}{*}NetPositivity_{jt}$$

Hard news precision (**H2**):

$$+\ \gamma_5\, DWidth_{jt} + \gamma_6\, DWidth_{jt}{*}NetPositivity_{jt} + \gamma_7\, Round_{jt} + \gamma_8\, Round_{jt}{*}NetPositivity_{jt}$$

Hyperbole (**H3**):

$$+\ \gamma_9\, |NetPositivity_{jt}| + \gamma_{10}\, |NetPositivity_{jt}|{*}NetPositivity_{jt}$$

Other: $+\ \gamma_{11}\, Litigation_{jt} + \gamma_{12}\, Litigation_{jt}{*}NetPositivity_{jt} + \gamma_{13}\, ForecastReputation_{jt}$

$$+\ \gamma_{14}\, ForecastReputation_{jt}{*}NetPositivity_{jt} + \gamma_{15}\, Size_{jt} + \gamma_{16}\, Size_{jt}{*}NetPositivity_{jt}$$

$$+\ \gamma_{17}\, Horizon_{jt} + \gamma_{18}\, Horizon_{jt}{*}NetPositivity_{jt} + \gamma_{19}\, AnalystFollowing_{jt}$$

$$+\ \gamma_{20}\, AnalystFollowing_{jt}{*}NetPositivity_{jt} + \gamma_{21}\, Certainty_{jt} + \gamma_{22}\, Certainty_{jt}{*}NetPositivity_{jt}$$

$$+\ \gamma_{23}\, PostFD_{jt}{*}NetPositivity_{jt} + \gamma_{24}\, PostSOX_{jt}{*}NetPositivity_{jt}$$

Controls: $+ \gamma_{25}\, FSURP_{jt} + \gamma_{26}\, OtherFSURP_{jt} + \gamma_{27}\, DWidth_{jt}{*}FSURP_{jt} + \gamma_{28}\, Round_{jt}{*}FSURP_{jt}$

$$+ \gamma_{29}\, Litigation_{jt}{*}FSURP_{jt} + \gamma_{30}\, ForecastReputation_{jt}{*}FSURP_{jt} + \gamma_{31}\, Size_{jt}{*}FSURP_{jt}$$

$$+ \gamma_{32}\, Horizon_{jt}{*}FSURP_{jt} + \gamma_{33}\, AnalystFollowing_{jt}{*}FSURP_{jt} + \gamma_{34}\, PostFD_{jt}{*}FSURP_{jt}$$

$$+ \gamma_{35}\, PostSOX_{jt}{*}FSURP_{jt} + \gamma_{36}\, |FSURP_{jt}| + \gamma_{37}\, |FSURP_{jt}|{*}FSURP_{jt} + \varepsilon_{jt}. \quad (2)$$

4.3.1 Coefficient predictions based on H1–H3

SignsAgree equals 1 if the signs of *FSURP* and *NetPositivity* agree and 0 otherwise. If management forecasts serve as contemporaneous verifiers of linguistic tone (H1a), then the coefficients on *SignsAgree+/+*NetPositivity* (good hard news verifies net positive tone) and *SignsAgree−/−*NetPositivity* (bad hard news verifies net negative tone) will be positive (γ_3, $\gamma_4 > 0$). If verification of net positive linguistic tone is more important to investors (H1b), then we expect the coefficient on *SignsAgree+/+* to be greater than coefficient on *SignsAgree−/−* ($\gamma_3 > \gamma_4$).

The hard news precision hypothesis (H2) predicts attenuation of price reaction to linguistic tone when forecasts are imprecise and thus more difficult to verify. *DWidth* equals 1 when *Width* is greater than the sample median and 0 otherwise. *Width* is the high minus low endpoints of a management range forecast, divided by price. Point forecasts are set to *Width* = 0, and minimum and maximum forecasts are set to the highest value of *Width* in the sample. *Round* equals 1 if the management forecast is perfectly divisible by $0.05 and 0 otherwise. H2 predicts negative coefficients on the *DWidth*NetPositivity* and *Round*NetPositivity* interactions (γ_6, $\gamma_8 < 0$).

If linguistic tone suffers credibility problems due to the hyperbole bias (H3), then the coefficient on the product of the absolute value of linguistic tone and linguistic tone, |*NetPositivity*|**NetPositivity*, will be negative ($\gamma_{10} < 0$), which indicates that the mapping of linguistic tone into returns is weaker for extreme tone.

4.3.2 Other potential cross-sectional determinants of the pricing of language

The model also includes other potential cross-sectional differences in the pricing of linguistic tone. Some of these effects have been examined in prior studies of the pricing of language but not in the context of a voluntary management forecast setting. Others have not been examined with respect to the pricing of language but have been found to be cross-sectional determinants of the pricing of earnings. To the extent that language predicts earnings, they are potential cross-sectional determinants of the pricing of language as well. At a minimum, these effects can be viewed as controls in the testing of our main hypotheses. To the extent that the effects have been found to exist for language in other contexts, our results provide evidence of the generalizability of the effect to the voluntary management forecast setting. To the extent that the effects have been documented for earnings but not for language, our results provide an incremental understanding of the pricing of language in the cross-section beyond our main hypotheses.

Litigation equals the probability of litigation estimated as the fitted value in the model provided by Rogers and Stocken (2005). Although managers have incentives to increase stock prices with optimistic disclosures, fear of litigation based on voluntary disclosures that are ex post over-optimistic tempers the tendency towards over-optimism in voluntary disclosures (Skinner 1994; Kasznik and Lev 1995; Baginski et al. 2002; Rogers and Stocken 2005). Rogers et al. (2011) document that managers' use of optimistic language increases litigation risk by showing that plaintiffs target optimistic statements in their lawsuits and that, controlling for a firm's economic conditions, sued firms have unusually linguistically optimistic earnings announcements. Accordingly, we expect investors will expect managers to avoid increased legal costs by tempering their optimism. The discipline provided by legal liability and the cost associated with legal exposure create more credible language, resulting in a greater price response to linguistic tone, suggesting that the coefficient on the *Litigation*NetPositivity* will be positive ($\gamma_{12} > 0$).

ForecastReputation equals the log of the count of the total number of management forecasts issued over the prior five years (maximum of one per period) times the average management usefulness of those forecasts computed as in Williams (1996).[17] Average forecast usefulness is zero if the firm did not issue a management forecast in the five prior years. Thus the proxy is increasing in the historical ability of managers to bring analysts closer to actual earnings and the length of the forecasting history available for investors to assess that ability. If investors believe that management forecast reputation is generalizable to other types of signals, then the coefficient on *ForecastReputation*NetPositivity* will be positive ($\gamma_{14} > 0$).

Prior literature documents an effect of the information environment on price reactions to disclosures in general. In a rich pre-disclosure information environment, information is impounded into price before news events. Using the typical proxy for richness of the pre-disclosure information environment, firm size, Atiase (1985) and Freeman (1987) empirically document a smaller price reaction to earnings news for larger firms. *Size* is the log of the firm's market value at the beginning of the forecast year. We predict a similarly attenuated response to the language accompanying the managerial forecasts of large firms ($\gamma_{16} < 0$).

Timing is a unique characteristic of the management forecast setting that also affects the richness of the information environment. Management forecasts are issued over different horizons. All else held equal, longer horizons are characterized by less competing information. As a result, earlier voluntary disclosures are likely to be a more important source of information for the resolution of uncertainty about the firm's future performance. *Horizon* equals the number of calendar days between the management forecast and the subsequent earnings release. We therefore expect that linguistic tone issued with longer horizon forecasts will be more informative in price setting ($\gamma_{18} > 0$).

Finally, Demers and Vega (2014) find that linguistic uncertainty and analyst following modify the pricing of linguistic tone in earnings releases. *Certainty* is a measure of linguistic certainty derived from the language in the management forecast press release using the Loughran and McDonald (2011)

---

[17] Williams (1996) defines usefulness as the absolute leading consensus analyst error minus the absolute management forecast error deflated by price. Positive values indicate that the analyst would become more accurate by simply parroting the management forecast. Combining usefulness and frequency is motivated by the empirical work of Hutton and Stocken (2009).

approach (multiplied by −100). *AnalystFollowing* is the log of the number of analysts providing forecasts at the management earnings forecast date.

4.3.3 Other control variables

We also include interactions of *NetPositivity* with two indicator variables to control for regulatory change. *PostFD* equals 1 for each management forecast observation issue after passage of Regulation Fair Disclosure and 0 otherwise. *PostSOX* equals 1 for each management forecast observation issue after passage of Sarbanes-Oxley and 0 otherwise. Finally, we control for the other information in the release (e.g., *FSURP, OtherFSURP*), and we interact *FSURP* with a set of modifiers identified in prior literature (e.g. forecast form, rounding, firm size, etc.), including its own absolute value to capture the S-curve effect in Freeman and Tse (1989).

4.4 Main results

Table 5 presents Pearson product-moment correlations among the variables used in the various regressions reported in this section. Although most of the correlations do not significantly differ from zero, some correlations are significant. Accordingly, we examine multi-collinearity diagnostics when estimating all regressions.

We first present tests of the individual effects of each hypothesized contemporaneous verifier in Table 6. Each column estimates an extended version of equation (1) that excludes control variables (other than *FSURP* and *OtherFSURP*, which appear in equation 1) and includes only one contemporaneous verifier. For obvious reasons, we do not base our final conclusions on the Table 6 results. However, we feel that it is useful to see the simple "univariate" effects of each variable on the pricing of language. For econometric reasons, we include intercept shifts on the effects of interest, but the verification effects are measured by the slope shift coefficients (i.e., the interactions) in each regression.

The results are consistent with our hypotheses. The first column shows that the effect of language on security prices is significantly stronger when the accompanying quantitative forecast sign "agrees" with the sign of the linguistic tone, regardless of whether the confirmed tone is positive or negative (H1a). Of particular interest is the coefficient estimate for *NetPositivity* in the first column. This coefficient captures the pricing of linguistic tone when its sign does not agree with the sign of *FSURP*. The coefficient is not significantly different from zero. The effect is stronger for agreement of two positive signs than for agreement of two negative signs, as indicated by the significant F-test on the equality of the two slope-shift coefficients (H1b).[18] The results in the second and third columns are consistent with the notion that the pricing of language accompanying less precise (and thus more difficult to verify) quantitative forecasts is attenuated (H2). The fourth column shows that extreme linguistic tone is also discounted (H3).[19]

Table 7 presents the results of the full model with all controls. Confirming our earlier findings, we find significant positive coefficients on the *FSURP* control variable and *NetPositivity*. Turning attention to the coefficients of interest on the slope shifts, all of our hypotheses are supported by the data. Significant positive (negative) slope shift coefficients indicate a stronger (weaker) pricing of linguistic tone. As predicted by H1a, the management forecast serves as a contemporaneous verifier of linguistic tone. The coefficients on the interactions between sign agreement and linguistic tone (*SignsAgree+/+*NetPositivity* and *SignsAgree−/−*NetPositivity*) are both significantly positive, consistent with the prediction that the agreement of the signs of the quantitative hard management forecast news and linguistic tone strengthen the price reaction to linguistic tone. The slope shift coefficient involving *SignsAgree+/+* (positive tone confirmed by hard news) is more than four times

[18] An alternative approach is simply to interact *NetPositivity* and *FSURP*. This interaction also captures the idea that the price reaction to *NetPositivity* depends on *FSURP*. However, it differs from our measure on two dimensions. First, it does not partition on sign. If the two variables are not symmetrically distributed around zero, then the interaction might indicate that, for example, small negative *FSURP* confirms small positive *NetPositivity* or small positive *FSURP* confirms small negative *NetPositivity*. H1 is about tone (positive versus negative), and thus we do not expect conflicting signs to increase stock price reaction to language. Second, the interaction would capture extreme magnitudes, which would confound it with our test of the hyperbole hypothesis (H3). In an untabulated analysis, we replaced the sign interactions with this sign and magnitude interaction. It is significant (t = 3.06), indicating that the price reaction to language depends on the level of the hard forecast news. We note, however, that this alternative untabulated regression has an R^2 of 15.1%, compared to 17.6% for the sign agreement only specification in Table 6.

[19] The size of the coefficient on the interaction of extreme linguistic tone with linguistic tone is driven simply by the scaling of the variables.

larger than the slope shift coefficient involving *SignsAgree−/−* (negative tone confirmed by hard news) and statistically significant based on the F-test of the difference between the coefficients reported at the bottom of the table. This finding is consistent with the prediction that management incentive-induced language inflation results in a stronger confirming effect of hard forecast news on positive linguistic tone (H1b).[20]

As predicted by H2, the coefficients on *DWidth*NetPositivity* and *Round*NetPositivity* are both significantly negative. The deterioration of quantitative forecast news precision, evidenced by wider forecast ranges and rounding, attenuates the pricing of linguistic tone. As predicted by H3, the coefficient on *|NetPositivity|*NetPositivity* is significantly negative, an indication that language hyperbole is discounted.

Aside from the results on our main hypotheses, we detect that litigation risk results in more credible linguistic tone with stronger associated price impacts, as evidenced by the weakly significant positive coefficient on *Litigation*NetPositivity*.[21] As documented by Demers and Vega (2014) in the earnings release setting, the information environment has a significant effect on the pricing of linguistic tone. The rich pre-disclosure information environment proxied by firm size attenuates the price response to linguistic tone, as evidenced by the significant negative coefficient on *Size*NetPositivity*. The existence of information intermediaries who produce competing information and demand high quality management disclosure strengthens the price response to linguistic tone, as evidenced by the significantly positive coefficient on *AnalystFollowing*NetPositivity*.[22]

---

[20] We re-estimated the model with alternative clustering approaches including by firm, by date, two-way clustering by firm and date, and two-way clustering by firm and quarter. All of our conclusions are unaffected except for one. In some of the alternative clustering approaches, the marginally significant coefficient on *SignsAgree−/−*NetPositivity* becomes marginally insignificant. However, as H1b hypothesizes, we expect the weakest result on this variable. Also, because we predict that the more precise signal (the management earnings forecast) verifies the less precise signal (language), and not the reverse, we do not have interactions of the sign agreement variables with *FSURP* in our model. If we include them, consistent with our expectations, they are insignificant, and our other results are not affected.

[21] A concurrent working paper by Bonsall et al. (2012) finds a similar result for the litigation effect in earnings releases.

[22] We do not find an effect for the *Certainty* modifier in contrast to Demers and Vega (2014). The nonsignificance may stem from the fundamental difference in our setting relative to their earnings release setting. In the management

In summary, our results from univariate and multiple regression tests confirm our predictions that characteristics of the quantitative management forecast news and language obtained from the management forecast announcement explain the price reaction to language.[23]

5 Additional tests

5.1 Post-announcement drift

Market efficiency is a maintained assumption in our analysis. However, in a descriptive vein, we replicate the Table 3 analysis to examine whether the incremental pricing of linguistic tone is delayed by replacing the dependent variable *AR* with *AR60*, an analogous 60-day post-announcement drift return accumulated beginning with the second day after the forecast release. We estimate three versions of the analogous pooled cross-sectional, time series, ordinary least squares model (with clustering and year effects as before):

$$AR60_{jt} = \lambda_0 + \lambda_1\, FSURP_{jt} + \lambda_2\, NetPositivity_{jt} + \lambda_3\, OtherFSURP_{jt} + \varepsilon_{jt} \qquad (3)$$

$$AR60_{jt} = \lambda_0^a + \lambda_1^a\, FSURP_{jt} + \lambda_2^a\, NetPositivity_{jt} + \lambda_3^a\, OtherFSURP_{jt} + \lambda_4^a\, Vfactor_{jt} + \lambda_5^a\, Vfactor_{jt}{*}NetPositivity_{jt} + \varepsilon_{jt} \qquad (3a)$$

$$AR60_{jt} = \lambda_0^b + \lambda_1^b\, FSURP_{jt} + \lambda_2^b\, NetPositivity_{jt} + \lambda_3^b\, OtherFSURP_{jt} + \lambda_4^b\, SignsAgree_{jt} + \lambda_5^b\, SignsAgree_{jt}{*}NetPositivity_{jt} + \varepsilon_{jt}, \qquad (3b)$$

where *Vfactor* equals the sum of four dichotomous verification proxies used to test our three hypotheses: *SignsAgree* (which equals one when the hard news verifies the sign of the language), minus a

forecast press release setting, the management forecast provides a contemporaneous verifier of linguistic tone, and the forecast has its own unique properties: width, rounding, forecast reputation, horizon, etc. In this setting, it appears that these unique properties dominate.

[23] In additional tests on reduced samples, we added additional controls for distress and industry concentration. The controls were intercept shifts and slope shift interactions of the variables with *NetPositivity.* These additional effects were insignificant and our main results were unaffected. We also redefined the announcement as the unit of analysis rather than the management forecast by discarding the annual forecast in announcements with both quarterly and annual forecasts (while still controlling for the information in the annual forecast through *OtherFSURP*). Our results on the variables of interest were not affected. Finally, we controlled for the data source of the announcement (i.e., Dow Jones Newswire versus PR Newswire) with additional intercept and slope shift coefficients with no effect on our results.

dichotomous variable equal to one if |*NetPositivity*| is above its sample median and zero otherwise (which measures the lower verifiability due to the hyperbole effect) minus *Dwidth* (which equals one when the hard forecast is less precise and thus less useful in verification) minus *Rounded* (which equals one when the hard forecast is less precise and thus less useful in verification).[24]

Coefficients λ_2, λ_2^a, and $\lambda_2^b < 0$ capture whether a pricing reversal occurs in the post-announcement period for *NetPositivity*. Equation (3) estimates whether the reversal exists. Equation (3a) estimates whether a reversal exists for conditions where the initial announcement period pricing was greatest (i.e., *Vfactor* is larger). Equation (3b) estimates whether a reversal exists for a single condition for which the announcement period pricing was greatest (i.e. when the signs of hard news and linguistic tone agree). Examining the sign agreement condition in isolation is motivated by the results in Table 6. Recall that *NetPositivity* was incrementally insignificant in the announcement period when signs did not agree.

Table 8 reports the estimation results. The first column shows the unconditional regression (equation 3). The coefficient on *NetPositivity* is not significantly less than zero, indicating that the pricing effects in the announcement period do not reverse in the 60-day post-announcement period. In fact, like Demers and Vega (2014), we detect a directionally consistent and statistically significant price drift associated with linguistic tone. The results persist when we condition on *Vfactor* (second column) and when we condition on *SignAgreement* (third column). Consistent with our results in Table 6 for announcement period returns, the association of linguistic tone with post-announcement returns also exists only for the case in which the sign of the hard news and linguistic tone agree.[25]

---

[24] For example, the highest value of *Vfactor* equals 1 when conditions exist that maximize contemporaneous verification. This occurs when signs of the hard forecast and linguistic tone agree, the forecast is not rounded, the forecast range is smaller, and the linguistic tone is not extreme.

[25] The coefficient on *NetPositivity* is negative but insignificant. Huang et al. (2014) document a price reversal on a proxy for abnormal tone in the context of earnings releases (i.e., a significant negative coefficient). To investigate whether our "unverified" tone would yield similar results as their abnormal tone measure, we repeated our estimation of equation (3″) on a smaller sample for which we could obtain data on a majority of control variables they use in their study (i.e., size, book-to-market, earnings volatility, return volatility, return level, and accruals). Our coefficient on *NetPositivity* did not achieve significance in the negative direction in this alternative estimation using either raw or ranked (as they did) values of the tone measure. We measure tone that is unverified by the hard

5.2 Attributions

Baginski et al. (2004) show that external attributions are associated with an increase in the association of unexpected earnings conveyed by a quantitative management forecast and stock returns. Therefore external attributions are a potential correlated omitted variable in our study if they are positively associated with any of our slope shift variables for which we expect a positive slope shift or negatively associated with any of our slope shift variables for which we expect a negative slope shift. For example, we hypothesize in H1 that hard news and linguistic tone sign agreement increases the strength of the relationship between net optimism and share price, a positive slope shift. Therefore, if the existence of an external attribution is positively correlated with sign agreement, the potential correlated omitted variable problem can exist. If external attributions are not correlated or negatively correlated with sign agreement, then omitting an external attribution by net optimism interaction in our model does not bias the coefficient on our hypothesized sign agreement by net optimism interaction.[26] As another example, we hypothesize in H2 and H3 that larger absolute values of net optimism, rounding, and range width decrease the strength of the relationship between net optimism and share price, a negative slope shift. Therefore, if external attributions are negatively correlated with absolute net optimism, rounding, or range width, the potential correlated omitted variable problem exists. External attributions that are not correlated or positively correlated with these variables do not cause a coefficient bias problem.

From our original sample (i.e., the sample before we read the text to identify missed bundling and other potential issues), we selected and read 244 documents to identify external attributions.[27] Using the classification method proposed by Baginski et al. (2004), we identified 244 external attributions (an average of 1 per document). We estimated a logistic regression model with a dichotomous dependent

---

information while they measure abnormal tone using an expected tone model. Furthermore, they investigate the drift after earnings releases and our sample excludes earnings releases.

[26] Our predictions are directional. The bias we refer to is a bias that would lead to an incorrect inference. In this example, a negative correlation is a bias in favor of a null hypothesis that we reject. That is, a negative correlation would work against our prediction and thus is a concern only if we were unable to reject the null.

[27] We randomly chose our sample from a set of observations that had not been subject to our data requirements and then applied necessary filtering procedures to arrive at a subsample of 244 that met all of our sampling criteria.

variable for external attributions (equal to one if present and zero otherwise; results not tabulated). The full set of main effects for hypothesized variables and controls (i.e., sign agreement, width, rounding, absolute net optimism, hard forecast surprise, and so on) were included as independent variables to explain external attributions, consistent with the exploratory approach in Baginski et al. (2004).

Absolute net optimism is significantly positively associated with the existence of an external attribution, which is not an issue because absolute net optimism is hypothesized in H3 to *decrease* the relation between language and stock prices. The fact that external attributions are positively associated with absolute net optimism works against our prediction. Forecast reputation and litigation risk are both negatively associated with external attributions, again not an issue because these variables are predicted to *increase* the relation between language and stock prices (and they are control rather than test variables in our main regressions).

Although this is a small sample, we re-estimated the results in Table 3. Then, we re-estimated again adding intercept and slope shifts for external attributions (i.e., external attribution dummy variable and an external attribution by net positivity interaction). The intercept and slope coefficients for external attributions are insignificant and the coefficient on net positivity remains significantly positive.

5.3 Other price-relevant content in the announcement

We exclude announcements containing earnings releases, and we rule out external attributions as correlated, omitted variables. In addition to earnings releases and external attributions, management forecasts can contain other content that could represent correlated, omitted variables. We performed two additional tests to rule out the effects of the other information. First, using the random sample described in Section 5.2, we read each article to directly identify other content. Only 65 of the 300 announcements (21.7%) contain additional content. The two most frequently occurring other items are sales reports (21) and restructuring or job cut announcements (12). We also identified instances of external financing-related news, such as stock repurchases and bond issuance or redemption (five), mentions of fraud or restatements (four), executive-related news such as new CEOs or CEO illness (four), litigation news such as the settlement of a legal case or expected litigation costs (three), production news such as opening a new plant or a new product approval (two), a strategic plan disclosure (two), and finally, a major customer that cut an order quantity (one). Using the same logistic regression methodology described in Section 5.2 with a dichotomous dependent variable for the existence of this other content, we found no evidence of an association between our main effects and the existence of the other content (results not tabulated). Using the same Table 3 re-estimation procedure described in Section 5.2, appending an intercept shift for the existence of special items and an interaction between special items and net positivity did not change our conclusions (results not tabulated).

We also perform an indirect test using customized tone measures. We generate our own dictionaries based on the word lists from the L&M dictionary and check to see whether our tone measure is mainly affected by the type of words used in the announcement: content-based words or framing-based words. It is difficult to argue that a particular word is content or framing based. However, nouns are more likely to be associated with a real economic event. A noun is self-explainable; for example, the L&M dictionary includes (among many) the following negative nouns beginning with the letters A and B: ABANDONMENT, ABUSE, ACCIDENT, ACCUSATION, ACQUITTAL, ADVERSARY, ALLEGATION, ALLEGATIONS, ARREARS, ASSAULT, ATTRITION, BAILOUT, BANKRUPTCY, BREACH, BREAKDOWN, and BRIBE, and positive nouns: ABUNDANCE, ACCOMPLISHMENT, ACHIEVEMENT, ADVANCEMENT, ADVANTAGE, ALLIANCE, ATTAINMENT, ATTRACTIVENESS, BENEFIT, BOOM, and BREAKTHROUGH.[28]

On the other hand, if the word is a verb, adjective, or adverb, it may be more likely to give action to, modify, or frame an event. That is, the word has less economic content in and of itself. Examples of negative words classified as verbs, adverbs, or adjectives include ABANDON, ABERRANT, ABNORMAL, ABOLISH, ABRUPT, ABUSE, ACCIDENTAL, ACCUSE, ADVERSARIAL,

[28] We reclassified all 2,349 negative words and 354 positive words from the L&M dictionary as a noun, verb, adverb, or adjective using the first listed part of speech provided by the Merriam-Webster Dictionary.

ADVERSE, AGAINST, ALLEGED, ARGUE, BARRED, and BREAK, and positive words: ABLE, ABUNDANT, ACCLAIMED, ACCOMPLISH, ADEQUATELY, ASSURE, ATTRACTIVE, BEAUTIFUL, BENEFICIAL, BETTER, BOOMING, and BRILLIANT.

We twice re-estimated our main model in Table 7, using a net optimism measure relying only on verbs, adverbs, and adjectives, once without control and once with control for a net optimism measure relying only on nouns. Our results on the hypothesized variables do not change in either specification (results not tabulated). Interestingly, the coefficient on the net optimism measure relying only on nouns is *less* significant that the measure relying on verbs, adverbs, and adjectives (although still significantly positive, t = 2.66). Although our dictionary reclassification method is indirect and not a perfect control for other economic content, combined with the direct analysis of other information in this section and with the knowledge that we discard observations bundled with earnings announcements leads us to conclude that other information in the management forecast announcement is not driving our results on the hypothesized variables.

5.4 Larger sample including bundled forecasts

As a final test to examine generalizability, we reintroduce bundled forecasts into the sample and control for the unexpected actual earnings. This increases our sample to 3,539 forecasts. The results for the test variables of interest are once again consistent with our main findings (results not tabulated).

6 Conclusion

Linguistic tone is relatively costless to provide and difficult to verify. In spite of these cheap talk characteristics, tone is incrementally priced to accompanying hard earnings news that is more easily verified. Using unbundled management earnings forecasts, we document a mechanism that serves as a timely alternative to ex post verification by which investors appear to assess the credibility of tone. We empirically link variation in the characteristics of quantitative disclosures that accompany language and a characteristic of the language itself to variation in the market pricing of linguistic tone. Specifically, when the signs of the quantitative management forecast and linguistic tone agree, tone's incremental pricing is strengthened. The effect is stronger when a good-news quantitative forecast confirms positive linguistic tone, consistent with the notion that optimistic language is less credible than pessimistic language due to management incentives to increase stock prices. Lower quantitative forecast quality and a characteristic of the tone itself, management's use of hyperbole for emphasis, attenuate the pricing of linguistic tone.

Language appears to be a powerful contributor to the pricing of disclosures in which it appears. Our findings increase the understanding of the pricing of language by identifying unique conditions that explain cross-sectional variation in that pricing. The majority of prior research has concentrated on the pricing of language in several different disclosure contexts, but very little research has considered the determinants of variation in the pricing of language, either in the cross-section or over time.

An interesting question is raised by the fact that we find linguistic tone to be incrementally priced in a sample of explicit forecasts of future earnings by management. The primary underlying economic reason to expect linguistic tone to be informative for prices is its predictive power for future earnings. Why then does language remain statistically and economically significant when it accompanies management-issued quantitative earnings *forecasts*? Managers may strategically use quantitative forecasts and linguistic tone to convey their expectations. If so, then some of the interesting work on the efficiency of management forecasts can be extended by considering language's incremental role in conveying expectations. For example, prior research has tested the efficiency of management forecasting by calculating the association of public information such as past accruals (Gong et al. 2000; Xu 2010) and past management forecast errors (Xu 2009) with subsequent management forecast accuracy. Future research might explicitly model the role of linguistic tone in communicating earnings expectations and whether the assessed efficiency of management is affected by consideration of the entire disclosure message.

References

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

Admati, A. R., & Pflederer, P. (2004). Broadcasting opinions with an overconfident sender. *International Economic Review* 45, 467-498.

Ajinkya, B., & Gift, M. (1984). Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22, 425–444.

Anderson, E. (1998). Customer satisfaction and word of mouth. *Journal of Service Research*, 5–17.

Anilowski, C., Feng, M. & Skinner, D. (2007). Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics* 44, 36–63.

Atiase, R. (1985). Predisclosure information, firm capitalization, and security price behavior around earnings announcements. *Journal of Accounting Research* 23, 21–36.

Baginski, S., Conrad, E. & Hassell, J. (1993). The effects of management forecast precision on equity pricing and on the assessment of earnings uncertainty. *The Accounting Review* 68, 913–927.

Baginski, S., & Hassell, J. (1990). The market interpretation of management earnings forecasts as a predictor of subsequent financial analyst forecast revisions. *The Accounting Review* 65, 175–190.

Baginski, S., & Hassell, J. (1997). Determinants of management forecast precision. *The Accounting Review* 72, 303–312.

Baginski, S., Hassell, J. & Kimbrough, M. (2002). The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in U.S. and Canadian markets. *The Accounting Review* 77 (1), 25–50.

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

Balakrishnan, K., & Bartov, E. (2010). Analysts' use of qualitative earnings information: Evidence from the IPO prospectus' risk factors section. Working paper, New York University.

Bamber, L., & Cheon, S. (1998). Discretionary management earnings forecast disclosures: Antecedents and outcomes associated with forecast venue and forecast specificity choices. *Journal of Accounting Research* 36, 167–190.

Bamber, L., Hui, K. & Yeung, E. (2010). Managers' EPS forecasts: Nickeling and diming the market? *The Accounting Review* 85, 63–95.

Benabou, R., & Laroque, G. (1992). Using privileged information to manipulate markets: Insiders, gurus, and credibility. *Quarterly Journal of Economics* 107, 921–958.

Bernard, V. (1987). Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (1), 1–48.

Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E. & Zinman, J. (2010). What's advertising content worth? Evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics* 125 (1), 263–305.

Blanes i Vidal, J. (2003). Credibility and cheap talk of security analysts: Theory and evidence. Working paper, London School of Economics, *SSRN eLibrary.*

Bonsall, S., Bozanic, Z., & Fischer, P. (2012). When is soft talk informative? Working paper, Penn State University, The Ohio State University, and the University of Pennsylvania, *SSRN eLibrary*.

Chen, J., Demers, E., & Lev, B. (2015). Oh what a beautiful morning! Diurnal variations in the tone of conference call communications. Working paper, State University of New York–Buffalo, University of Virginia - Darden, and New York University.

Chuk, E., Matsumoto, D. & Miller, G. (2013). Assessing methods of identifying management forecasts: CIG vs. Researcher collected. *Journal of Accounting & Economics* 55, 23–42.

Crawford, V., & Sobel, J. (1982). Strategic information transmission. *Econometrica* 50, 1431–1451.

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

Davis, A., & Tama-Sweet, I. (2012). Managers' use of language across alternative disclosure outlets: Earnings press releases versus MD&A. *Contemporary Accounting Research* 29, 804–837.

Dambra, M., Wasley, C., & Wu, J. (2013). Soft-talk management cash flow forecasts: Bias, quality, and stock price effects. *Contemporary Accounting Research* 30, 607–644.

Demers, E., & Vega, C. (2011). Linguistic tone in earnings press releases: News or noise? Working paper, University of Virginia – Darden and the Federal Reserve Board of Governors, *SSRN eLibrary.*

Demers, E., & Vega, C. (2014). The Impact of Credibility on the Pricing of Managerial Textual Content. Working paper, University of Virginia – Darden and the Federal Reserve Board of Governors, *SSRN eLibrary*.

Dye, R., & Sridhar, S. (2004). Reliability-relevance trade-offs and the efficiency of aggregation. *Journal of Accounting Research* 42, 51–88.

Feldman, R., Govindaraj, S., Livnat, J. & Segal, B. (2010). Management tone change, post earnings announcement drift and accruals. *Review of Accounting Studies* 15, 915–953.

Fischer, P., & Stocken, P. (2010). Analyst information acquisition and communication. *The Accounting Review* 85, 1985–2010.

Freeman, R. (1987). The association between accounting earnings and security returns for large and small firms. *Journal of Accounting & Economics* 9, 195–228.

Gong, G., Li, L. & Xie, H. (2009). The association between management earnings forecast errors and accruals. *The Accounting Review* 84, 497–530.

Gow, I., Taylor, D. & Ormazabal, G. (2010). Correction for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85, 483–572.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review* 95 (1), 384–394.

Hafzalla, N. (2009). Managerial incentives for discretionary disclosure: evidence from management leveraged buyouts. *Review of Accounting Studies* 14 (4), 507–533.

Huang, X, Teoh, S. & Zhang, Y. (2014). Tone management. *The Accounting Review* 89, 1083–1113.

Hurkens, S., & Kartik, N. (2009). Would I lie to you? On social preferences and lying aversion. *Experimental Economics* (June), 180-192.

Hutton, A., & Stocken, P. (2009). Prior forecasting accuracy and investor reaction to management earnings forecasts. Working paper, Boston College and Dartmouth College. *SSRN eLibrary.*

Hutton, A., Miller, G., & Skinner, D. (2003). The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41, 867–890.

Jennings, R. (1987). Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. *Journal of Accounting Research* 25, 90–110.

Kartik, N. (2009). Strategic communication with lying costs. *Review of Economic Studies* 76, 1359–1395.

Kasznik, R., & Lev, B. (1995). To warn or not to warn: Management disclosure in the face of an earnings surprise. *The Accounting Review* 70 (1), 113-134.

King, R., Pownall, G. & Waymire, G. (1990). Expectations adjustments via timely management forecasts: Review, synthesis, and suggestions for future research. *Journal of Accounting Literature* 9, 113–114.

Kothari, S.P., Shu, S., & Wysocki, P. (2009). Do managers withhold bad news? *Journal of Accounting Research* 47, 241–276.

Kross, W., Ro, B., & Suk, I. (2011). Consistency in meeting or beating expectations and management earnings forecasts. *Journal of Accounting and Economics* 51, 37–57.

Li, F. (2010). The information content of forward-looking statements in corporate filings – a naïve Bayesian machine learning approach. *Journal of Accounting Research* 48, 1049–1102.

Loughran, T., & &, B. (2011). When is a liability not a liability? *Journal of Finance* 66, 35–55.

Mangen, C., & Durnev, A. (2011). The real effects of disclosure tone: Evidence from restatements. *SSRN eLibrary.*

Merkley, K., & Bamber, L. & Christensen, T. (2013). Detailed management earnings forecasts: do analysts listen? *Review of Accounting Studies* 18, 479–521.

Michels, J. (2012). Do unverifiable disclosures matter? Evidence from peer-to-peer lending. *The Accounting Review* 87, 1385–1413.

Morgan, J., & Stocken, P. (2003). An analysis of stock recommendations. *Rand Journal of Economics* 34, 183–203.

Mullainathan, S., Schwartzstein, J. & Shleifer, A. (2008). Coarse thinking and persuasion. *The Quarterly Journal of Economics* 123 (2), 577–619.

Ng, J., Tuna, A., & Verdi, R. (2013). Management forecast credibility and underreaction to news. *Review of Accounting Studies* 18, 956–986.

Peterson, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies* 22, 435–480.

Rogers, J., & Van Buskirk, A. (2012). Bundled forecasts in empirical accounting research. *Journal of Accounting and Economics* 55 (1), 43–65.

Rogers, J., Van Buskirk, A. & Zechman, S. (2011). Disclosure tone and shareholder litigation. *The Accounting Review* 86 (6), 2155-2183.

Rogers, J., & Stocken, P. (2005). Credibility of management forecasts. *The Accounting Review* 80, 1233–1260.

Skinnner, D. (1994). Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (1), 38–60.

Stocken, P. (2000). Credibility of voluntary disclosure. *The RAND Journal of Economics* 32 (2), 359–374.

Waymire, G. (1984). Additional evidence on the information content on management earnings forecasts. *Journal of Accounting Research* 22, 703–718.

Williams, P. (1986). The relation between a prior earnings forecast by management and analyst response to a current management forecast. *The Accounting Review* 71, 103–116.

Xu, W. (2009). Evidence that management earnings forecasts do not fully incorporate information in prior forecast errors. *Journal of Business, Finance, and Accounting* 36, 822–837.

Xu, W. (2010). Do management earnings forecast incorporate information in accruals? *Journal of Accounting and Economics* 49, 227–246.

Table 1 Sample selection

Initial sample of management earnings forecasts from Dow Jones Business News and Press Releases Newswire (1997–2006)	6,180
Lacking required First Call coverage or other required data	(2,136)
Bundled with earnings releases[a]	(2,035)
Article contains multiple firms or discussion of price reactions[b]	(245)
Final sample of unbundled management earnings forecasts	1,764
Distinct firms	750

[a]We read the text of forecasts initially identified as unbundled to verify that we have removed any forecasts bundled with earnings releases. The discarded forecast observations include those identified by our reading of the text.

[b] Identified by reading the disclosure text.

Table 2 Descriptive statistics for 1,754 management earnings forecasts

Variable	Mean	Standard deviation	25th percentile	Median	75th percentile
Panel A: Variable distributions for continuous variables					
AR	-0.053	0.128	-0.106	-0.028	0.023
NetPositivity	-0.003	0.011	-0.010	-0.002	0.004
FSURP	-0.002	0.033	-0.005	-0.000	0.003
Size	7.225	1.750	6.009	7.027	8.347
ForecastReputation	-0.003	0.019	-0.001	0.000	0.001
Horizon	121.4	138.9	25.0	49.0	183
Width	0.021	0.058	0.000	0.001	0.003
Certainty	18.30	14.56	9.12	14.15	22.00
Litigation	0.218	0.337	0.004	0.031	0.289
AnalystFollowing	2.002	0.795	1.609	2.079	2.564

Panel B: Variable distributions for dichotomous categorical variables

Variable	Yes = 1 (%)	No = 0 (%)
SignsAgree	1,165 (66.4%)	589 (33.6%)
Round	446 (25.4%)	1,308 (74.6%)
PostFD	1,076 (61.3%)	678 (38.7%)
PostSOX	610 (34.8%)	1,144 (65.2%)

Panel C: Form distribution

Variable	Number of observations (%)
Point	356 (20.3%)
Range	1,232 (70.2%)
Minimum	94 (5.4%)
Maximum	72 (4.1%)

Panel D: Year Distribution

Year	Number of observations (%)
1997	97 (5.5%)
1998	196 (11.2%)
1999	220 (12.5%)
2000	253 (14.4%)
2001	282 (16.1%)
2002	229 (13.1%)
2003	163 (9.3%)
2004	116 (6.6%)
2005	112 (6.4%)
2006	86 (4.9%)

AR is the size and book-to-market adjusted three-day announcement period return. *NetPositivity* is the language metric derived from the management forecast as described in the text using the Loughran and McDonald (2010; L&M) financial dictionaries. *FSURP* is the management forecast surprise measured as the point (or midpoint of the range or disclosed minimum or maximum) forecast minus the preceding median consensus analyst forecast, scaled by price. *Size* is the log of the firm's market value at the beginning of the forecast period.

Table 2 continued

ForecastReputation equals the log of the count of the total number of management forecasts issued over the prior five years (maximum of one per period) times the average management forecast usefulness computed as in Williams (1996). It is equal to zero if the firm did not issue a management forecast in the five prior years. *Horizon* equals the number of calendar days between the management forecast and subsequent earnings release. *Width* equals the high minus low endpoints of the management range forecast, divided by price. (Point forecasts are set to *Width* = 0; minimum and maximum forecasts are set to the highest value of *Width* in the sample.) *Certainty* is a measure of linguistic certainty derived from the language in the management forecast press release using the L&M dictionaries. *Litigation* equals probability of litigation estimated as the fitted value in the model proposed by Rogers and Stocken (2005). *AnalystFollowing* is the log of the number of analysts providing forecasts at the management earnings forecast date. *SignsAgree* equals 1 if the signs of *FSURP* and *NetPositivity* agree and 0 otherwise. *Round* equals 1 if the management forecast is perfectly divisible by $0.05 and 0 otherwise. *PostFD* equals 1 for each management forecast observation issue after passage of Regulation Fair Disclosure and 0 otherwise. *PostSOX* equals 1 for each management forecast observation issue after passage of Sarbanes-Oxley and 0 otherwise.

Table 3 Is the linguistic tone in management forecasts incrementally priced? (*AR* is the dependent variable.)

Variable	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
Intercept	-0.047 (-4.06)***	-0.028 (-2.45)**	-0.027 (-2.37)**
FSURP	0.578 (3.88)***		0.506 (3.36)***
NetPositivity		3.287 (8.62)***	3.058 (8.01)***
OtherFSURP			-0.092 (-1.72)
N	1,754	1,754	1,754
R^2	7.6%	13.2%	14.5%
Firm/Date Clustered, Year Effects	Yes	Yes	Yes

***/**/* indicate significance at the 0.01/0.05/0.10 levels in a one-tailed test (two-tailed for intercept).

Good (bad) news is defined as *FSURP* ≥ 0 (*FSURP* < 0), where *FSURP* is the management earnings forecast surprise as defined in Table 1.

Positive (Negative) tone is defined as *NetPositivity* ≥ 0 (*NetPositivity* < 0), where *NetPositivity* is the linguistic tone as defined in Table 1.

Table 4 Are hard management forecast news and linguistic tone associated?

Panel A: Sign agreement

	Sign of Linguistic Tone	
	Positive (%)	Negative (%)
Hard Forecast News:		
Good News (%)	642 (36.6%)	308 (17.6%)
Bad News (%)	281 (16.0%)	523 (29.8%)
χ^2 test of independence	= 185.9***	

Panel B: Correlation

Pearson product moment correlation between *FSURP* and *NetPositivity*	0.175 ***
Spearman rank order correlation between *FSURP* and *NetPositivity*	0.337 ***

Good (bad) news is defined as *FSURP* ≥ 0 (*FSURP* < 0), where *FSURP* is the management earnings forecast surprise as defined in Table 1.

Positive (Negative) tone is defined as *NetPositivity* ≥ 0 (*NetPositivity* < 0), where *NetPositivity* is the linguistic tone as defined in Table 1.

***/**/* indicate significance at the 0.01/0.05/0.10 levels in a one-tailed test.

Table 5 Correlations among regression variables

		1)	2)	3)	4)	5)	6)	7)	8)	9)	10)	11)	12)	13)	14)	15)	16)
1)	*AR*																
2)	*NetPositivity*	0.33*															
3)	*SignsAgree*	0.02	-0.07*														
4)	*DWidth*	-0.04	-0.03	-0.04													
5)	*Round*	0.00	-0.02	-0.02	0.15*												
6)	$\lvert NetPositivity \rvert$	-0.13*	-0.41*	0.16*	0.03	0.05											
7)	*Litigation*	-0.03	-0.03	0.02	-0.06*	-0.02	-0.01										
8)	*Size*	0.09*	-0.05	0.00	-0.21*	0.04	-0.00	-0.01									
9)	*Horizon*	-0.00	0.08*	-0.05	0.19*	0.26*	-0.03	0.01	0.10*								
10)	*ForecastReputation*	-0.04	-0.06	0.02	-0.06*	-0.10*	0.03	0.03	-0.19*	-0.15*							
11)	*AnalystFollowing*	0.05	-0.03	0.00	-0.13*	0.00	-0.05	0.15*	0.59*	0.02	-0.09*						
12)	*Certainty*	0.02	0.15*	-0.05	-0.02	-0.02	-0.22*	-0.02	0.06*	0.13*	-0.04	0.10*					
13)	*PostFD*	0.17*	0.19*	-0.03	0.02	0.03	-0.20*	0.14*	0.22*	0.13*	-0.03	0.16*	0.26*				
14)	*PostSOX*	0.14*	0.19*	-0.04	0.02	0.13	-0.19*	0.14*	0.14*	0.10*	-0.02	0.25*	0.29*	0.57*			
15)	*FSURP*	0.15*	0.17*	-0.01	-0.05	-0.00	-0.19*	-0.02	0.02	-0.00	0.17*	0.00	0.02	0.02	0.04		
16)	*OtherFSURP*	0.03	0.05	0.04	-0.01	-0.00	-0.02	0.01	-0.04	-0.02	0.33*	-0.00	-0.01	-0.01	-0.01	0.55*	
17)	*AbsFSURP*	-0.02	-0.03	-0.01	0.17*	0.13*	0.06*	-0.01	-0.08*	0.19*	-0.41*	-0.10*	0.00	-0.05	-0.07*	-0.62*	-0.50*

* indicates significant Pearson correlation at the 0.01 level. See Table 1 for variable definitions.

Table 6 Univariate tests of the market verification of linguistic tone in management earnings forecasts for individual concurrent verifiers

Variable	Expected sign	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
Intercept	None	-0.049 (-3.86)***	-0.023 (-1.95)*	-0.027 (-2.21)**	-0.027 (-2.08)**
NetPositivity	+	0.038 (0.06)	3.613 (7.20)***	3.448 (9.02)***	7.205 (7.71)***
FSURP``	+	0.283 (1.92)**	0.519 (3.43)***		
OtherFSURP	+	-0.069 (-1.45)	-0.093 (-1.68)		
Agreement hypothesis (**H1a**):					
SignsAgree	None	0.019 (2.03)**			
SignsAgree+/+ NetPositivity*	+	5.604 (5.52)***			
SignsAgree-/- NetPositivity*	+	3.438 (3.77)***			
Management incentives hypothesis (**H1b**): F-statistic on difference between slope coefficients		4.75**			
Hard news precision hypothesis (**H2**):					
DWidth	None		-0.009 (-1.44)		
DWidth NetPositivity*	-		-1.113 (-1.82)**		
Round	None			-0.003 (-0.46)	
Round NetPositivity*	-			-1.458 (-2.00)**	
Hyperbole hypothesis (**H3**):					
|*NetPositivity*|	None				-0.323 (-0.50)
|*NetPositivity*|**NetPositivity*	-				-217.48 (-4.64)***
R^2		17.6%	14.8%	14.8%	16.7%
N		1,754	1,754	1,754	1,754

***/**/* indicate significance at the 0.01/0.05/0.10 levels in a two-tailed test (one-tailed for specific sign predictions). See Table 1 for variable definitions. Estimated using firm/date clustering and year fixed effects.

Table 7 Market verification of linguistic tone in management earnings forecasts (multiple regression analysis)

Variable	Expected sign	Coefficient	t-statistic
Intercept	None	-0.045	-1.97**
NetPositivity	+	7.719	4.09***
SignsAgree	None	0.009	0.87
SignsAgree+/+ NetPositivity* **(H1a)**	+	6.533	4.48***
SignsAgree-/- NetPositivity* **(H1a)**	+	1.414	1.29*
DWidth	None	0.001	0.23
DWidth NetPositivity* **(H2)**	-	-1.359	-2.36***
Round	None	0.003	0.44
Round NetPositivity* **(H2)**	-	-1.391	-2.36***
|*NetPositivity*|	None	-2.604	-2.58**
|*NetPositivity*|**NetPositivity* **(H3)**	-	-229.764	-4.94***
Additional potential cross-sectional determinants:			
Litigation	None	-0.013	-1.06
Litigation NetPositivity*	+	1.867	1.57*
ForecastReputation	None	0.211	1.24
ForecastReputation NetPositivity*	+	16.925	1.10
Size	None	0.003	1.28
Size NetPositivity*	-	-0.670	-3.23***
Horizon	None	-0.000	-1.47
Horizon NetPositivity*	+	0.002	0.95
AnalystFollowing	None	0.001	0.34
*AnalystFollowing*NetPositivity*	+	0.891	2.09**
Certainty	None	-0.001	-2.56**
*Certainty*NetPositivity*	+	0.008	0.23
*PostFD*NetPositivity*	None	0.359	0.38
*PostSOX*NetPositivity*	None	0.014	0.01
Additional Controls:			
FSURP	+	3.315	4.70***
OtherFSURP	+	0.068	1.10
DWidth FSURP*	-	-0.323	-1.67**
Round FSURP*	-	-0.429	-1.97**
Litigation FSURP*	+	0.070	0.21
*ForecastReputation*FSURP*	+	-1.920	-0.70
*Size*FSURP*	-	-0.322	-3.65***
*Horizon*FSURP*	+	0.001	1.37*
*AnalystFollowing*FSURP*	+	0.267	1.44*
*PostFD*FSURP*	None	0.014	0.05
*PostSOX*FSURP*	None	0.413	1.35
|*FSURP*|	None	-0.187	-0.78
|*FSURP*|**FSURP*	-	-8.319	-3.54***
R^2		27.1%	
N		1,754	
H1b: F-stat on difference between **H1a** coefficients	8.18***		

***/**/* indicate significance at the 0.01/0.05/0.10 levels in a two-tailed test (one-tailed for specific sign predictions). See Table 1 for variable definitions. Estimated using firm/date clustering and year fixed effects.

Table 8 Post-management forecast drift (*AR60* is the dependent variable.)

Variable	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
Intercept	-0.007 (-0.47)	-0.009 (-0.59)	-0.017 (-0.96)
FSURP	0.346 (1.72)*	0.341 (1.69)*	0.200 (0.96)
NetPositivity	1.074 (2.06)**	1.720 (2.34)**	-0.820 (-0.84)
OtherFSURP	0.006 (0.10)	0.010 (0.15)	0.022 (0.33)
Vfactor		-0.002 (-0.47)	
*Vfactor*NetPositivity*		0.589 (1.39)	
SignsAgree			0.012 (1.10)
*SignsAgree*NetPositivity*			2.647 (2.55)**
N	1,754	1,754	1,754
R^2	2.9%	3.0%	3.4%
Firm/Date Clustered, Year Effects	Yes	Yes	Yes

***/**/* indicate significance at the 0.01/0.05/0.10 levels in a two-tailed test.

See Table 1 for variable definitions.

AR60 is the size and book-to-market adjusted 60-day post-announcement period return.

Vfactor equals the sum of four dichotomous verification proxies: *SignsAgree* (which equals 1 when the hard news verifies the sign of the language), minus a dichotomous variable equal to 1 if |*NetPositivity*| is above its sample median and 0 otherwise (which measures the lower verifiability due to the hyperbole effect) minus *Dwidth* (which equals 1 when the hard forecast is less precise and thus less useful in verification) minus *Rounded* (which equals 1 when the hard forecast is less precise and thus less useful in verification).