Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 951

October 2008

Soft Information in Earnings Announcements: News or Noise?

Elizabeth Demers

and

Clara Vega

Soft Information in Earnings Announcements: News or Noise?*

Elizabeth Demers
Accounting & Control Area
INSEAD
Boulevard de Constance
77305 Fountainebleau Cedex
France
liz.demers@insead.edu

Clara Vega
Board of Governors
of the Federal Reserve System
International Finance Division
20th Street & Constitution Avenue, NW
Washington, DC 20551
clara.vega@frb.gov

Abstract

This paper examines whether the "soft" information contained in the text of management's quarterly earnings press releases is incrementally informative over the company's reported "hard" earnings news. We use Diction, a textual-analysis program, to extract various dimensions of managerial net optimism from more than 20,000 corporate earnings announcements over the period 1998 to 2006 and document that unanticipated net optimism in managers' language affects announcement period abnormal returns and predicts post-earnings announcement drift. We find that it takes longer for the market to understand the implications of soft information than those of hard information. We also find that the market response varies by firm size, turnover, media and analyst coverage, and the extent to which the standard accounting model captures the underlying economics of the firm. We also show that the second moment of soft information, the level of certainty in the text, is an important determinant of contemporaneous idiosyncratic volatility, and it predicts future idiosyncratic volatility.

JEL Classifications: G14; D82; M41
Keywords: Soft information; earnings announcements; post-earnings drift; cheap talk; earnings quality; information uncertainty; momentum; voluntary disclosure

*We thank workshop participants at Duke University, the Swiss Finance Institute in Lugano, and the National University of Singapore, in addition to Per Olsson, Katherine Schipper, Steve Sharpe, Jeremy Stein, and especially Jennifer Francis for valuable comments and discussions. We also appreciate the excellent research assistance of Armel Allard, Shannon Demers, Nicholas Klagge, Marcus Newman, and Hien Tran, we thank Craig Carroll for processing our text files using the Diction software, and we thank PR Newswire for providing us with the earnings announcement data. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or any other person associated with the Federal Reserve System.

1. Introduction

Beginning with Ball and Brown (1968), many researchers have examined stock price responses to corporate earnings announcements. An extensive subsequent literature investigates the market responses to other elements of firms' voluntary disclosures issued simultaneously with the earnings news. Most of these prior studies have examined market responses to quantitative, *hard*, largely verifiable information disclosed by management. In this study, we consider the firm's abnormal stock market price and volatility responses to managers' *soft* information disclosures during quarterly earnings announcements, incremental to the impact of the hard earnings surprise.[1] Specifically, we use a well-established linguistic algorithm to extract two dimensions of managerial soft information – net optimism and certainty - from over 20,000 corporate earnings announcements filed with the PR Newswire service during the period of January 1998 through July 2006. In terms of the first dimension, we find that the unexpected component of manager's net optimism is significantly associated with short-window announcement period returns, and that it also predicts post-earnings announcement drift. However, we find that it takes relatively longer for soft information to be incorporated into asset prices than for hard information. Further tests show that the pricing of this form of soft information depends upon whether conditions are present to induce the provision of informative soft information rather than noise, and it also depends upon the informativeness of the simultaneously released hard earnings data. In terms of the second dimension, we find that managerial certainty is inversely associated with increased idiosyncratic volatility during the short-window announcement interval, and that it also predicts abnormal idiosyncratic volatility during the intermediate-term post-announcement period.

The process of price discovery in financial markets remains poorly understood. Numerous studies show that quantitative information about fundamentals explain only a small portion of asset price movements and volatility.[2] Thus

[1] Petersen (2004) defines hard information to be data that is quantitative, easy to store and transmit in impersonal ways, and whose content is independent of the collection process. By contrast, soft information is characterized as that which is directly verifiable only by the person who collected and produced it, that cannot be unambiguously documented (Stein (2002); Petersen (2004)), and that is often communicated in text (Petersen (2004)). The concept of soft information is thus closely related to that of "cheap talk," where the latter may be characterized as any claim that is costless, unbinding, and non-verifiable (Krishna and Morgan (2008)).

[2] See Cenesizoglu and Timmermann (2008) for a recent review of the literatures related to the forecasting of the first and second moment of stock returns, which can be briefly summarized as showing that fundamentals explain a very small portion of asset price movements.

identifying additional sources of fundamental information that are incorporated into asset prices is of basic importance to financial economics. One such source is the qualitative verbal information managers communicate in quarterly earnings announcements (henceforth, "soft information"). A key feature of this type of information is that, in contrast to the hard earnings figure which is verifiable and thus has the potential to convey useful information (Grossman (1981) and Milgrom (1981)), theory does not always predict that soft information will convey valuable information (e.g., Crawford and Sobel (1982) and Benabou and Laroque (1992), amongst others). Furthermore, an extensive prior literature documents strategic managerial behavior in the derivation of quarterly earnings (Burgstahler and Dichev (1997); Degeorge, Patel and Zeckhauser (1999)), and in the release and/or presentation of other quantitative, and thus much less subtle, information at the time of earnings announcements (e.g., Schrand and Walther (2000); Bradshaw and Sloan (2002); and Bhattacharya, Black, Christensen and Larson (2003)). Hence, the *a priori* information content of management's soft linguistic information is far from obvious. We find, however, that managers' unexpected net optimism does have information content incremental to the simultaneously released hard earnings news. We further find that the level of certainty expressed in managers' language is inversely associated with the firm's abnormal return volatility. Our findings of significant market reactions to soft information disclosures are consistent with the view that repeated interactions between managers and investors may be sufficient to support truthful revelation of the managers' soft information, perhaps because the informed agent's possible current gains from opportunistic behavior can be wiped out by future losses in payoff from damaged reputation (see, e.g., Sobel (1985) and Stocken (2000)).[3]

In extended tests we find that firms' net optimism is differentially incorporated into asset prices. First, we document that net optimism is priced more for high tech firms, firms with high PE ratios, and companies with lower quality accounting data, suggesting that the role of soft information in the price formation process is a function of the characteristics of the competing hard data available to market participants. Second, we find that net optimism is priced more for stocks with

[3] Our findings are also consistent with the recent surge in the use of algorithms that are designed to automatically read and code economic data releases for the purpose of generating trading orders that can be executed even more quickly than human analysts are able to finish reading the first line of the press release (The Economist, June 21, 2007) and thus with the associated rise in the number of firms that sell such algorithms (e.g., Ravenpack at http://www.ravenpack.com/index.html).

greater analyst following and higher levels of media coverage. The latter results are consistent with the notion that the market response to net optimism is a function of the credibility of the soft information, where credibility is assumed to be enhanced by repeated two-way interactions between management and information intermediaries.[4] Third, we find that the price responsiveness to net optimism is increasing in the stock's turnover. Consistent with prior literature, we view turnover as a proxy for the level of disagreement among informed traders about the value of the firm, and thus we interpret our results as suggesting that a lack of consensus over the value of the firm generates a demand for managers' soft information.[5] An alternative interpretation is that low turnover stocks are also stocks that are harder to short-sell and short-selling constraints reduce the adjustment speed of prices to new information (Diamond and Verrecchia (1987)). Inconsistent with this latter explanation, we find that net optimism of low turnover firms is not differentially associated with future short- and medium-term abnormal returns relative to that of high turnover firms (i.e., there is no evidence of a greater delayed price response to soft information for less actively traded stocks). Thus, we conclude that the heightened response of high turnover firms to net optimism during the announcement interval is not directly driven by the speed of adjustment to this new information, but rather by the demand for soft information under conditions of high dispersion in informed investors' beliefs.

Our study extends the soft information literature by examining the role of another linguistic measure extracted from managerial press releases, certainty, in explaining announcement period and post-announcement idiosyncratic volatility. We find that the level of certainty expressed in managers' earnings announcements is inversely related to idiosyncratic volatility during the announcement period, and that it is also a leading indicator for post-announcement abnormal volatility. The results for the association between uncertainty in managers' language and the second moment of stock returns are robust to controlling for fundamental measures of uncertainty in the firm's economic environment.

Our study relates most closely to the soft information studies of Tetlock, Saar-Tsechansky and Macskassy (2008), Engelberg (2007), Li (2006), and Davis, Piger

[4] Krishna and Morgan (2004) show that repeated two-way communication, such as that suggested by the interaction of analysts and the media with management, will help to improve the informational content of soft information conveyed by the agent, even if the analysts and media are uninformed.

[5] The demand for information is highest when the asset payoff variance is high (see, Grossman and Stiglitz (1980); Veldkamp (2006)) and dispersion of beliefs are high (Foster and Viswanathan (1996)).

and Sedor (2007). Tetlock, et al. (2008) and Engelberg (2007) examine whether a quantitative measure of negative language in firm-specific earnings news stories can be used to predict firms' future accounting earnings and stock returns. They conclude that linguistic media content captures otherwise hard to quantify aspects of firms' fundamentals which investors quickly incorporate into stock prices. Our study differs from, and is complementary to, both studies in several respects. First, they examine the association between *media-expressed* negativity and future measures of firm performance, while we examine the relation between *management-expressed* net optimism and both contemporaneous and future stock returns and idiosyncratic volatility. Relative to the media, managers have different insights, motivations, biases, and fiduciary duties in their communications to parties who are external to the firm. This potential misalignment of interests, together with the subtlety of the language constructs derived from press releases, enable us to test cheap talk theories in the context of managerial earnings announcements. Second, Tetlock, et al. (2008) and Engelberg (2007) examine one dimension of language, negativity, while we consider the role of both unexpected net optimism as well as certainty in explaining asset price dynamics.[6] Third, while Tetlock, et al. (2008) have a longer time series of observations for a sample that is restricted to very large, highly transparent S&P 500 firms, we have a shorter, more recent time series of observations that span a much broader sample of firms that are not all subject to high information environments such as those in the S&P 500. The advantage of this broader sample in the context of our study is that it provides the cross-sectional variation necessary to examine whether and how firm characteristics affect the market's response to soft information.

Davis et al. (2007) also explore the association between management-expressed optimism and pessimism in earnings announcements and market returns, however their analysis is restricted to the mean effect of soft information during the 3-day announcement period. Our study extends their analyses along two dimensions.

[6] Li (2006) also investigates the notion of certainty but does so using a very different linguistic approach from ours; he adopts a count of a few researcher-specified "risk" and "uncertainty" words included in corporate annual reports whereas our study uses a sophisticated, externally validated linguistic algorithm to extract management expressed certainty from quarterly corporate earnings announcements. He examines the relation between his measure of risk sentiment and both future earnings and stock returns. In contrast to Li (2006) who finds that risk sentiment predicts future returns in a cross-sectional setting, we find a relation between the certainty in manager's earnings releases and stock prices only in the short-window announcement period, not in the post-announcement longer-term window. We document a relation between our measure of certainty and abnormal idiosyncratic volatility during the announcement window, and we also show that certainty is a predictor of post-announcement idiosyncratic volatility.

First, we extend their short-window announcement abnormal returns tests in several ways: i) by documenting that unexpected managerial net optimism also predicts post-earnings announcement drift; ii) by documenting that the market's response to net optimism is an increasing function of media and analyst coverage (mechanisms that enable repeated two-way communications as predicted by theoretical cheap talk models); and iii) by showing that the market's response to soft information is increasing in the relative lack of informativeness of the alternative, hard earnings information, such as for high-tech firms that have complex business models for which the standard accounting model less aptly captures the underlying economics of the firm. Second, using the same Diction linguistic software adopted by Davis et al. (2007), we extract a third dimension of management soft information, certainty, and explore its relation to idiosyncratic volatility. We document that the use of more wavering (i.e., less direct) language in managerial earnings announcements is associated with abnormal idiosyncratic stock volatility both during the 3-day announcement window and during the longer post-announcement drift period, after controlling for fundamental measures of economic uncertainty.

The rest of this paper is organized as follows. Section 2 describes our samples, data sources, and the measurement of our soft information variables. Section 3 examines the relation between the unexpected component of managerial net optimism and, respectively, short-window announcement period returns and the post-announcement drift phenomenon. In Section 4 we explore the relation between managerial certainty and idiosyncratic stock price volatilities, while Section 5 provides a summary and conclusion to our study.

2. Sample and Data Description

2.1 Samples

We obtain the text of quarterly earnings announcements for the period of January 1998 through July 2006 from PR Newswire. We are able to match using the ticker symbol and the announcement date (allowing for a 3-day window discrepancy) for 27,705 of the PR Newswire observations with the CRSP/Compustat database (4,771 different firms) and 17,484 of these are matched to IBES (3,372 different firms). Hereafter we refer to these two samples as the "Compustat" and "IBES" samples, respectively. We include only those observations for which we can calculate earnings surprises, 3-day abnormal returns surrounding the earnings announcement, and 60-trading-day abnormal returns both prior to, and post-, announcement. We also drop observations with stock prices below $1 and above $10,000 and firms with negative book values. After imposing all of the preceding restrictions, we are left with a final sample of 3,764 firms (2,610 firms) and 21,580 firm-quarter (13,907 firm-quarter) observations for the Compustat (IBES) sample. In untabulated results we find that the firms in our Compustat sample report slightly higher earnings surprises, have slightly higher ROA, and are larger (on the basis of total sales, market capitalization and total assets) than the firms in the CRSP-Compustat universe. They also have higher P/E ratios (but not market-to-book ratios) and are more likely to report special items such as impairments and restructuring charges than firms in the corresponding population. Our IBES sample firms have slightly lower market-to-book ratios and are slightly larger than the CRSP-Compustat-IBES universe of firms. The IBES sample firms are not significantly different from the corresponding population on the basis of profitability (ROA), earnings surprises, trading volume, or P/E ratios.

Throughout this study, we tabulate and discuss the results of all of our tests using each of the Compustat and IBES samples, respectively, and we do so for several reasons. First, the IBES constraint imposes a bias in favor of the inclusion of firms that are larger and subject to richer information environments, while we are also interested in understanding the role of soft information for the broader universe of firms that are not subject to such high exposure and associated analyst filtering mechanisms. Second, Graham, Harvey and Rajgopal (2006) report that 85.1% of CFO survey respondents considered earnings in the same quarter of the prior year to be the most important earnings benchmark, followed secondly by the analyst consensus estimate at 73.5%. The CFOs interviewed in their study further noted that the first item in their press release is often a comparison of the current quarter's earnings with four-quarters-lagged earnings. Accordingly, we expect that the prior year's same quarter actual earnings provides the framing context for management's current earnings announcement even if it is not the figure associated with the strongest market response for firms that are tracked by analysts.[7]

[7] Although a recent study by Ljungqvist, Malloy and Marston (2007) suggests that the currently available IBES data may be subject to non-random ex post changes, the concerns that those authors raise relate to analyst *recommendations* rather than the analyst *estimates* that we use in our study.

2.2 Data

We obtain market values, stock returns, and trading volume from the Center for Research in Security Prices (CRSP) databases. Historical accounting data are obtained from Compustat, while IBES provides the alternative source for historical earnings realizations that are matched to analyst estimates. We obtain media counts from the Factiva database.

Corporate quarterly earnings announcements are provided by PR Newswire, with each firm-quarter's announcement being furnished as an individual text file. Prior to subjecting these files to the linguistic algorithm processing described below, we undertake a number of analyses upon, and make a number of modifications to, the announcements. First, we use keyword searches to develop indicator variables for the presence of an income statement, a balance sheet, and a statement of cash flows, respectively, in each announcement file. Next, we identify tabulated figures in the text (including the financial statements) by searching for strings of numbers, and where identified we cut these elements from the files so that tables of figures are not confounding the textual linguistic analysis.[8] Third, using mechanical search algorithms that we designed based upon extensive manual review of the announcements, we separately remove the company description and "safe harbor" paragraphs from the announcements so that only the earnings announcements themselves remain in the text files to be analyzed.[9]

2.3 Measuring Net Optimism and Certainty

There has been an increased interest in recent years in determining the sentiment and degree of certainty conveyed in public communications by government institutions, the media, and corporate entities. Various methods have been employed to measure the soft information contained in these communications and to systematically analyze its impact on market measures of activity and individual

[8] The language algorithms typically count each numerical expression as a "word" and thus leaving numerical tables in the files will confound the measurement of the linguistic constructs that we wish to extract from the texts by exaggerating both the total number of words as well as the numerical term scores. We explicitly include other variables designed to capture the presence and/or contents of the quarterly financial statements.

[9] The company description sections typically describe the entity in extremely positive terms, whereas the safe harbor provisions include many uncertainty-related expressions. Thus, their inclusion would have the effect of increasing the net optimism, positivity, and uncertainty linguistic scores in an artificial manner in the sense that neither of these sections is directly related to the managerial earnings announcement news *per se* that we seek to analyze.

behavior. For example, Ehrmann and Fratzscher (2007) analyze the style of communication among central banks by manually classifying Reuters press releases in terms of economic outlook and policy inclinations. Lucca and Trebbi (2008) design an automated scoring method to measure the content of central bank communications about future policy rate moves and find that medium-term and long-term government bond yields react to their soft information measure but not to current policy rate decisions. Das, Martinez-Jerez and Tufano (2005) examine the connection between on-line discussion, news activity, and movement in stock prices by developing their own net optimism index based upon five distinct language processing algorithms that classify discussion as bullish, bearish, or neutral, while Das and Chen (2007) use the same method to extract small investor net optimism from stock message boards. Li (2006) uses a simple count of the relative frequency of "risk" and "uncertainty" in corporate annual reports and relates this to future earnings and stock returns. Numerous studies use Diction software to extract linguistic characteristics from various texts (e.g., Bligh and Hess (2007); Ober, Zhao, Davis and Alexander (1999); Yuthas, Rogers and Dillard (2002); and Davis, et al. (2007)),[10] while Tetlock (2007), Tetlock, et al. (2008), and Engelberg (2007) use *General Inquirer* ("GI"), an alternative linguistic algorithm, to measure the level of negativity in media content and relate this to securities returns.

In our primary reported tests, we use version 6.0 of the Diction text-analysis program to measure the level of optimism, pessimism, and certainty, respectively, in managers' earnings announcements. Diction is a well-established language processing algorithm that has been used extensively in prior research to analyze the

[10] Bligh and Hess (2007) use the Diction software to measure "certainty, pessimism, optimism, activity, immediacy and jargon" in 45 FOMC statements, 44 congressional testimonies and 105 speeches given by the Chairman of the Federal Reserve, Alan Greenspan, between May 18th, 1999 and June 30th, 2004. They conclude that Greenspan's rhetoric predicts movements both in the Treasury forward rates and in the federal funds future rates. Ober, Zhao, Davis and Alexander (1999) use Diction to assess corporations' use of certainty in public communications by examining the "Management Discussion and Analysis" (MD&A) section of the10-K reports of the six Fortune 500 companies with the largest increases in profits and the six companies with the largest decreases in profits in 1996 in each of six major industry groups. They find that corporations with large profit increases do not use rhetoric with significantly more certainty than companies that experience large decreases in their profits, and thus they conclude that managers "tell it like it is" and "avoid weasel words." Yuthas, Rogers and Dillard (2002) investigate the characteristics of corporate annual reports in order to ascertain whether corporate communication is ethical and conforms to Habermas' four principles of comprehensibility, truthfulness, sincerity, and legitimacy. They find that managers of firms with bad performance generally engage in "ethical discourse" by not distorting the truth about their companies' financial position. However, they also report that these firms strategically use "fewer self-referential terms," perhaps in order to disassociate bad performance with internal factors. Davis, Piger and Sedor (2007) examine the use of pessimistic and optimistic language in earnings press releases.

speeches of Federal Reserve policymakers, political speeches, corporate annual reports, and earnings announcements.[11] The algorithm uses a series of thirty-three dictionaries (word-lists) to search text passages for different semantic features such as, e.g., praise, satisfaction, or denial. For our study, we analyze the earnings announcements using the *optimism* and *certainty* definitions of Diction. *Optimism* is defined as, "language endorsing some person, group, concept or event or highlighting their positive entailments" while *certainty* is defined as, "language indicating resoluteness, inflexibility, and completeness and a tendency to speak *ex cathedra*" (Digitext Inc. (2000)). The Diction formula for *net optimism* is [praise + satisfaction + inspiration]-[blame + hardship + denial].[12] Following prior studies, we interpret the first and second components of the optimism formula as "*optimism*" and "*pessimism,*" respectively, and we refer to the difference between the two as "*net optimism.*" Similar to General Inquirer (GI), Diction generates a ratio representing the number of words in the target article that are contained in a particular word-list dictionary divided by the total number of words in the article multiplied by 100.[13,14] The word lists are mutually exclusive except for homographic terms. Using this Diction-based procedure generates measures of *optimism* and *pessimism* that are each bounded by 0 and 100.

The Diction formula for *certainty* is [tenacity + leveling + collectives + insistence] - [numerical terms + ambivalence + self reference + variety]. We redefine this formula to include numerical terms as additive to certainty rather than subtracting them from the score. In the context of earnings announcements, which may include both management's analyses of past results as well as their future expectations, we view the provision of more hard, ex post verifiable quantitative information to be

[11] See http://www.dictionsoftware.com/files/dictionresearch.pdf for a more extensive summary of published academic studies using the Diction software.

[12] The terms associated with each of the characteristics that generate the optimism and certainty variables are reproduced in Davis, et al. (2007) and are available in extended detail in Digitext Inc. (2000).

[13] Technically speaking the ratio in Diction is for every 500 words, so dividing the Diction metrics by 5 provides a measure that is directly comparable to the GI measures.

[14] The Diction program also allows the user to select different communication "norms" that generate linguistic sentiment scores based upon comparisons of the target text to Diction's database of 22,027 codified texts. The Diction texts range from, e.g., campaign speeches (2357 observations) to corporate financial reports (48 different texts) that originate from the period 1948 to 1998. Rather than adopt any such Diction normalizations that are based upon a very small sample size of financial texts that largely predate our sample period, we use only the simple ratios of dictionary words to total words in the text passages to generate our raw measures of optimism, pessimism, and certainty. Our approach is thus similar to the methodology underlying the General Inquirer program as well as to the Diction-based measures adopted by Davis et al (2007).

indicative of more direct and precise expression rather than the use of more obtuse language.[15] In order to obtain measures for certainty that are of comparable magnitudes to optimism and pessimism, we normalize the calculated variable by adding the absolute value of the lowest (i.e., negative) valued raw certainty score, dividing the sum through by the maximum value, and then multiplying by 100. Hence our *certainty* measure is also bounded by zero and 100.

In specification checks, we also use the General Inquirer (GI) program introduced into the finance literature by Tetlock (2007), Tetlock, et al. (2008), and Engelberg (2007). Specifically, we use GI to measure the *negativity (positivity)* in the earnings announcement text, which is simply defined as the percentage of negative (positive) words from the Harvard IV-4 psychological dictionary to the total words in the announcement text.[16] We calculate the difference between positivity and negativity and use this as our GI-based measure of net positivity. Standard dictionary definitions of the words "negativity" ("positivity") and "pessimism" ("optimism") are different, and consistent with this the dictionary list of words that GI and Diction, respectively, associate with each of these linguistic sentimental constructs is also different. As one would expect, however, GI's *negativity (positivity)* sentiment and Diction's *pessimism (optimism)* sentiment are correlated measures (ρ=0.40).[17]

---

[15]Diction's presumption is that "numerical terms hyper-specify a claim, thus detracting from its universality." This may be true in the context of political speeches and some other forms of expository prose that formed the original basis for Diction, but in extensive readings of earnings announcements we found that the more numerical terms included in the announcement, the closer was the soft information to hard (verifiable) information, and the less room there was for ambiguity. We also found that managers tend to quote fewer numbers (e.g., they are less likely to provide forecasts) when uncertainty is high, so that the number of numerical terms divided by the number of words in the announcement is negatively correlated with present and future stock return volatility. However, we find that the variable certainty is a better predictor of present and future stock return volatility than the simple ratio of the number of numerical terms divided by the number of words in the announcement, and hence Diction's certainty measure is indeed capturing aspects of the underlying constructs beyond just the greater precision provided by numbers versus prose. As previously noted, we calculate the number of numerical terms in the announcement after having excluded any income statements, balance sheets, and statements of cash flows provided in the earnings announcement, and we control for the existence and contents of the financial statements separately in the regressions.

[16] For more information on the GI program, the reader is referred to Stone, Dunphy, Smith and Ogilvie (1966), Tetlock (2007), or the GI website at: http://www.wjh.harvard.edu/~inquirer/.

[17] The main differences between GI and Diction are the word-lists and the handling of homographic terms (i.e., words with identical spelling but that denote different objects or activities, such as the word "state"). In the development of their algorithms, GI developers subjectively assess the most common usage of each homographic word, while the Diction program weights each such word according to the findings of Easton (1940). In our empirical tests we find that the word list used in GI to construct net optimism works better than the word list used in Diction in predicting current and future stock returns.

2.4 Measuring Hard and Soft Information Surprises

2.4.1 Earnings Announcement Surprises

We use two alternative measures of the firm's unexpected quarterly earnings. One measure uses the median IBES forecast and the other uses last year's same quarter earnings per share (i.e., a seasonal random walk model) as the market expectation of earnings. In particular, we define unexpected earnings as $UE_{jt} = A_{jt} - E_{jt}$, where A_{jt} is the announced earnings per share of firm *j* on day *t*, E_{jt} is either last year's same quarter earnings per share for the Compustat sample (E_{jt-4}) or the IBES median forecast for the IBES sample. To make a meaningful comparison of the estimated surprises across firms and measures, we follow the literature and use standardized surprises. Specifically, we divide the surprise by the firm-specific standard deviation of the forecast error, defining standardized unexpected earnings (hard information) associated with firm *j* at time *t* as

$$SUE_{jt} = \frac{A_{jt} - E_{jt}}{\hat{\sigma}_j},$$

where $\hat{\sigma}_j$ is the standard deviation of the forecast error, $A_{jt} - E_{jt}$, estimated using the entire Compustat and IBES sample of observations for each respective firm. We require each firm to have non-missing earnings data in the Compustat and IBES databases, respectively, for 10 quarters. Alternatively, to prevent a hindsight bias, we estimate the standard deviation of the forecast error using the firm's previous 20 quarters of unexpected earnings data following Bernard and Thomas (1989) and Tetlock et al. (2008). We also allow for a trend in Compustat unexpected earnings for all firms with more than four years of earnings data. Our regression results and SUE-sorted portfolio returns are qualitatively the same when we use these alternative measures and are available upon request.

2.4.2 Measuring Surprises in Net Optimism

Similar to the standard specification for hard earnings surprises, we adopt an expectations model for net optimism in order to attempt to capture the "surprise" element of the level of net optimism contained in management's press release. Only the unexpected component of net optimism should be reflected in the announcement period abnormal returns. Untabulated results show that the level of net optimism contained in management's most recent prior quarter's announcement is the best expectation for this quarter's net optimism, and accordingly we use a non-seasonally-adjusted random walk model to calculate the unexpected net optimism as $\Delta NetOpt_{jt}=NetOptimism_{jt}-NetOptimism_{jt-1}$.[18]

We similarly define the standardized unexpected net optimism associated with firm j at time t as follows,

$$SNetOpt_{jt} = \frac{\Delta NetOpt_{jt} - \mu_{\Delta NetOpt}}{\sigma_{\Delta NetOpt}},$$

where $\Delta NetOpt_{jt}$ is the difference between the net optimism of firm j in quarter t, estimated using either Diction or General Inquirer (GI), and the net optimism of the most recent prior quarter's announcement by firm j, $\mu_{\Delta NetOpt}$ and $\sigma_{\Delta NetOpt}$ are equal to the mean and standard deviation of $\Delta NetOpt$ across all firms and all quarters in our sample. Alternatively, to prevent a hindsight bias and prevent any biases induced by a trend in the data, we estimate the standard deviation of the forecast error using last quarter's standard deviation of the forecast error across firms and our results are qualitatively similar.[19] Because $\mu_{\Delta NetOpt}$ and $\sigma_{\Delta NetOpt}$ are constant for any firm j, the standardization will not affect either the statistical significance of the response estimates or the fit of the regression. The standardization of each of the hard earnings and soft net optimism surprise variables enables us to make meaningful comparisons of asset price responses to the two different news items.

[18] The adjusted R-squared of the seasonally-adjusted random walk model is 19.55% compared to 24.12% for the non-seasonally-adjusted random walk model. The Akaike and Schwarz information criteria as well as out-of-sample predictive tests also favor the latter model. This result is in contrast to the earnings per share model selection for which our results are consistent with prior studies; the seasonally-adjusted random walk model outperforms the non-seasonally-adjusted model (the adjusted R-squared of the former is 23.97% compared to 13.90% for the latter). Although using the surprise component in sentiment is the theoretically correct specification and is consistent with the earnings surprise specification, in untabulated results we also find that raw sentiment affects asset prices. The latter result is consistent with the notion that in the announcements managers compare their earnings performance with last year's performance, and hence the raw net optimism expressed is already implicitly relative to an expectational benchmark. The market's response to raw sentiment is, however, weaker than the response to the "surprise" sentiment, i.e., the market realizes that net optimism is serial correlated. We note that Engelberg (2007) and Tetlock (2007) and Tetlock et al. (2007) analyze raw sentiment in news articles, presumably because the media is reporting "news," and hence it may not be necessary in such media-based research settings to estimate the "surprise" in the media's sentiment.

[19] We do not allow the mean and standard deviation of unexpected sentiment to be different across firms in the standardization process because doing so would reduce our sample size considerably; our PR Newswire sample does not provide a sufficiently long time series for most of the firms for us to compute firm-specific standard deviations of the forecast error.

2.5 Descriptive Statistics

Table 1A provides descriptive statistics for the soft information variables for each of the two samples. Each of the soft information variables exhibits a considerable range of values in both samples. As shown, optimism has a mean value of about 1.3 for both samples, while pessimism has a mean value of about 0.6. In other words, about 1.3 out of every 100 words are found in Diction's "optimism" dictionary, which translates into approximately 11 optimism-increasing words, on average, per earnings announcement based upon mean word counts of 825 and 888 for each of the Compustat and IBES samples, respectively. On average, net optimism is slightly lower for the Compustat sample at 0.66 versus 0.71 for the IBES firms. Panels A and B of Appendix A present example texts from firms with relatively high and low optimism scores, respectively, with the words contributing to the high optimism and pessimism scores shown in highlights. The difference in economic circumstances and thus tone between the optimistic and pessimistic announcements is obvious, and the linguistic algorithm aptly captures this.

Relative to optimism and pessimism, the certainty variable exhibits considerably higher mean and median scores because of the relatively heavy use of numerical terms within the texts of earnings announcements. Excerpts from announcements scoring high and low on the certainty measure are reproduced panels C and D of Appendix A, respectively. As shown in Panel C, the tone of the Raytheon announcement is entirely assertive and almost every statement is supported by a numerical term or financial comparison, yielding Raytheon a relatively high certainty score. In contrast, the announcement of Play-By-Play Toys & Novelties makes use of relatively little precise numerical support for management's assertions and in addition uses imprecise language including such terms such "approximately," "uncertain," "expect," and "believe" rather than clear and strong assertions of fact.[20]

Table 1B presents the correlation matrix for the soft information variables. As shown, certainty is not highly correlated with any of the other soft information measures nor are optimism and pessimism highly correlated with one another or with the earnings surprise variable (SUE). This combination of results suggests that managers, on average, present a discussion in their earnings announcements that is

[20] Two of the components of the certainty score, insistence and variety, are based upon word repetitions and word variety rather than dictionary words per se. These linguistic constructs do not lend themselves to being highlighted in the announcement example texts.

directional (i.e., either optimistic or pessimistic) rather than balanced, and that the soft information of the press release conveys different information from that conveyed in the hard earnings surprise.

Tables 2A and 2B present descriptive statistics for the firms in our Compustat and IBES samples, respectively.[21] As shown, the quarterly earnings surprise is positive, on average, and larger for both samples when calculated using analyst earnings expectations (SUE_IBES) compared to the surprises generated using a Compustat-based seasonal random walk model (SUE). As expected, the IBES sample is constrained to relatively larger firms that, on average, are followed by more analysts, have higher levels of media coverage, higher turnover, and have higher earnings quality as captured by the lower value of the EFKOS e-loading factor.

3. The Relation Between Soft Information and Stock Returns

In this section we revisit the work of Tetlock, et al. (2008) and Engelberg (2007) using *managerial* earnings announcements rather than *media* stories. We also extend the basic results of Davis, et al. (2007), who document a mean earnings announcement period market response to net optimism, in several ways: i) by testing cheap talk theories and information-driven hypotheses to investigate whether and how firm characteristics affect the market's response to soft information; and ii) by investigating the association between soft information and the post-announcement drift phenomenon.

3.1 The Announcement Period Response to Net Optimism

We first investigate the announcement period response to the hard and soft information surprises contained in the earnings announcement. Our dependent variable is defined as the 3-day, size- and book-to-market-adjusted cumulative abnormal returns (CARs) for the period [-1, +1] where 0 is the earnings announcement day. Specifically, to calculate abnormal returns we subtract the contemporaneous returns on size- and B/M-matched portfolios. The portfolios are constructed using the method of Fama and French (1992). For June of the current year, we classify all firms into 25 portfolios by size at the end of June of the current year and by B/M at the end of December of the previous year. We only use stocks

[21] Detailed variable definitions are provided in Appendix B.

with positive book values (data item 60 on the Compustat tapes) to calculate size and B/M breakpoints. The resulting portfolios are then equally weighted.[22]

3.1.1 Baseline Announcement Period Pricing Tests

We first examine the relative market responses to standardized earnings and managerial net optimism surprises using a pooled regression model. We expect that the hard earnings news will have a larger effect on asset prices relative to the soft information because the earnings surprises are more uniform, verifiable, have a greater likelihood of being understood, and on an overall basis are likely to be more credible than the soft information.[23] We use the following pooled regression to examine whether $\beta_1 > \beta_2$:

$$\sum_{i=-1}^{1} AR_{jt+i} = \beta_0 + \beta_1 SUE_{jt} + \beta_2 SNetOpt_{jt} + \varepsilon_{jt}. \qquad (1)$$

The results of these tests are reported in Panel A of Table 3. The standard errors reported in all tables are clustered by firm and calendar quarter to allow for correlation in error terms across firms and quarters.[24] As shown, both the hard earnings surprise ("SUE") and the surprise component of net optimism ("SNetOpt") are statistically very significant, both across the Compustat and IBES samples as well as across the alternative GI- and Diction-based measures of net optimism. The adjusted R^2 and coefficients on the earnings surprise variables are generally similar to those reported in prior earnings response studies. Consistent with the results of Tetlock, et al. (2008) and Engelberg (2007) who find that the stock prices respond to *media* negativity, our results show that the market also considers the net optimism expressed by *management* in their earnings announcements to be credible and

[22] We adopt this methodology because Barber and Lyon (1997) and Daniel and Titman (1997) suggest that matching sample firms to firms of similar sizes and book-to-market ratios, rather than using factor betas, yields better-specified test statistics. For further details on this methodology please refer to Fama and French (1992).

[23] Although the quarterly earnings figures are not audited, the annual financial results that are the composite of the quarterlies are audited at the end of the firm's fiscal year. Furthermore, the financial results are all prepared in accordance with generally accepted accounting policies ("GAAP"). This standardization combined with the ex post audit of the annual figures enhances the understandability and credibility of the hard earnings measures reported by management.

[24] Our results are robust to using Newey-West and Panel Corrected Standard Errors (PCSE), however based upon the diagnostic tests suggested by Petersen (2008), the most appropriate standard errors for the model specifications in our study are those clustered by firm and calendar quarter.

informative.[25] Our results are consistent with Davis, et al. (2007) who find a similar announcement period response to optimism and pessimism, respectively, in managerial announcements. However, our use of standardized variables enables us to speak to the *relative* market response to hard and soft information as the coefficients on SUE and SNetOpt can be interpreted as the CAR response to a one standard deviation change in each of the hard and soft earnings surprises, respectively. We find, as expected, that the coefficient on the standardized earnings surprise variable is considerably larger than that on standardized net optimism surprise. Our findings suggest that although the market clearly impounds managerial net optimism quickly into prices, investors nevertheless weight the more objective, hard earnings information more heavily than the soft information during the announcement window.

The results of interactively adding *size* to the equation (1) regression are shown in Panel B of Table 3, with size defined as the natural log of the firm's market capitalization as of the last fiscal year end. As shown, the market response to both hard and soft information is lower for larger firms, consistent with the notion that these firms operate in richer information environments, with the result that both the soft and hard news embedded in the firms' earnings announcements are at least partially anticipated by market participants and thus generate a lower announcement period price response.

3.1.2 The Impact of Firm Characteristics on Announcement Period Response

Table 4 reports the results from a number of specifications that individually include in equation (1), in addition to the size-interacted terms, additional explanatory variables interacted with each of the hard earnings and soft information surprise variables. Because of the high level of correlation among some of the candidate independent variables, we first consider the impact of each variable separately. We then show the results from a single multivariate regression that includes all of the proxies that are not related to similar underlying constructs (and thus for which there

[25] Similar to Engelberg (2007), Tetlock et al. (2007), and Tetlock (2007), we find that GI's negativity measure affects asset prices more than GI's positivity measure. We also find that Diction's pessimism affects asset prices slightly more than optimism, but the difference is not as large as that found using the GI measures. All of our conclusions are robust to allowing negative and positive soft information surprises to differentially affect asset prices, so for the sake of parsimony we report only the results from symmetric specifications.

is no strong theoretical correlation).[26] In what follows, we discuss only the coefficients on the newly added interactive terms unless the results for the earnings surprise, soft information surprise, and size-interacted surprise variables are inconsistent with those reported for the baseline regressions.

Given the lack of both cost and (even *ex post*) verifiability associated with the soft linguistic information that is the subject of our study, we view this data as a form of "cheap talk." Hence, our initial analyses involve using our soft information variable, *SNetOpt*, to empirically test cheap talk theories. We first examine whether analyst and media coverage of the company respectively impact the intensity of the market's response to soft and hard information. Assuming that information conveyed by managers is credible, there is no *prima facie* reason to expect the asset prices of firms that are more widely covered by analysts and the media to react differently to information released during the announcement period.[27] However, theoretical models of cheap talk question the usefulness of this information (e.g., Crawford and Sobel (1982) and Benabou and Laroque (1992)), and the literature has proposed certain mechanisms to solve or at least alleviate the misinformation problem. One such mechanism is proposed by Krishna and Morgan (2004), who suggest that repeated two-way communication improves the informational content of cheap talk. Since the presence of journalists and analysts facilitates two-way communication, we expect that their presence would increase the credibility and usefulness of the soft information conveyed by managers. In Table 4 Panel A and B we show, respectively, the effect of analyst coverage, measured as the log of one plus the number of analysts covering the firm, and of media coverage, measured as the number of times a firm is mentioned in the headline or lead paragraph of an article from newswire services in

[26]For example, turnover, media exposure, analyst coverage, and institutional ownership are highly correlated variables that capture elements of similar underlying constructs. High tech firms, high PE ratio firms, firms with high R&D expenditures and firms with high EFKOS e-loadings are highly correlated as well.

[27] Previous literature finds analyst and media coverage to be important determinants of PEAD, but in general they do not affect the 3-day CAR reaction to news. One exception is Peress (2008), who finds that asset prices react more to the earnings announcement surprises of firms that have high media coverage during the announcement period. His explanation is that investors suffer from limited attention and do not react to the news of "neglected" firms. Because of our concern for simultaneity bias, our measure of media coverage does not include media mentions during the announcement period, so our results are not directly comparable to Peress (2008). Nevertheless, our findings are consistent with his in that we also document that asset prices react more to hard earnings surprises released by firms with high media coverage, however this result is only statistically significant for our Compustat sample where limited attention may be more important.

the previous 60 trading days before the earnings announcement date [t-62,t-2].[28] The interaction of analyst coverage with net optimism surprise is positive and significant in both the Compustat and IBES regressions, suggesting that firms that are more heavily followed have higher price responses to managerial net optimism incremental to the size effect. We interpret these results as supportive of the Krishna and Morgan (2004) theory, which is to say that under conditions of greater analyst scrutiny, together with the corresponding potential for two-way communication between the information intermediaries and managers, managers are induced to convey more truthful net optimism rather than noisy cheap talk.

An alternative explanation for our results is that firms with higher levels of analyst following are simply informationally more efficient (i.e., impound information more quickly into prices). For this interpretation to hold, however, a symmetrical result on the SUE term interacted with analyst coverage would also be expected. To the contrary, however, the SUE interacted term is insignificant for both samples. Thus, we conclude that the credibility of the net optimism conveyed by "neglected" firms is more questionable than that of heavily followed firms, resulting in neglected firms' soft information generating a lower market response. We find weaker evidence in favor of the Krishna and Morgan (2004) theory when we examine the interaction term between media coverage and net optimism. This interaction term is also positive, but it is only significant for the Compustat sample, and we find that the interaction between media coverage and SUE is also significant for the Compustat sample, which is consistent with the Peress (2008) finding related to limited attention biases.

We next examine whether *turnover*, measured as the average of the natural log of de-trended turnover, defined as de-trended daily volume of shares traded divided by stock outstanding, cumulated over the 60-trading-day pre-announcement period [-62, -2], has an impact on the market's response to managerial net optimism.[29] We predict that the price response to soft information will be increasing in the turnover of

[28] We use Factiva to extract this measure and only take into account publications that have over 500,000 current subscribers. The list of data sources is: The Wall Street Journal (all editions), Associated Press Newswire, the Chicago Tribune, the Globe and Mail, Gannett News Service, the Los Angeles Times, the New York Times, the Washington Post, USA Today and all Dow Jones newswires.

[29] We use the de-trended measure of turnover because turnover is not stationary. Following Campbell, Grossman and Wang (1993), we calculate the turnover trend as a rolling average of the past 60 trading day turnover. We add back the mean of turnover to our de-trended measure so that the units are economically meaningful.

the stock for at least two reasons. First, Chan (2003) documents that turnover and media coverage are highly correlated (and this result is corroborated in unreported tests for our sample), and we have previously documented a weakly greater response to soft information for firms with higher levels of media coverage. Second, turnover is commonly used in the empirical literature as a proxy for the dispersion in informed traders' beliefs, a tenet that is also supported by numerous theoretical models (e.g., Harris and Raviv (1993), Wang (1998), and Hong and Stein (2003)).[30] We posit that a lack of consensus amongst informed traders will lead managers to provide more useful soft information, because the demand for information is highest when dispersion of beliefs is high (Foster and Viswanathan (1996)) and when uncertainty is high (e.g., Grossman and Stiglitz (1980) and Veldkamp (2006)). Based upon this reasoning, we expect to find a positive coefficient on the *SNetOptXTurnover* variable. Consistent with the demand side hypothesis for price responsiveness to soft information, the results reported in Table 3 Panel C show a significant and positive coefficient on the *SNetOptXTurnover* variable for both samples.

An alternative interpretation for our turnover results is that firms with low turnover are hard to short-sell, and the short-sale constraint impedes investors from quickly and fully punishing managers for delivering uninformative soft information. This reasoning leads to the expectation that managers of short-sale constrained firms tend to provide only noisy cheap talk, resulting in the observed smaller price responsiveness to *SNetOpt* for lower turnover stocks. We consider this alternative explanation to be unlikely, however, because it requires that managers of such short-sale constrained firms be overly myopic since investors can punish them by selling the stock that they own, by shorting over time as liquidity becomes available, and by not providing any future financing. However, as a specification check we interact *SNetOpt* with institutional ownership, a commonly-used alternative proxy for short-sale constraints. In untabulated results using *SNetOpt* interacted with institutional ownership as an alternative proxy for short-sale liquidity, we again find evidence consistent with the demand-side hypothesis rather than the short-sale alternative as the institutional ownership interacted variable is not significant. As a further

[30] One exception in the literature is the model of Foster and Viswanathan (1996), which implies that there is a negative correlation due to a "waiting game" equilibrium. However, Kandel and Pearson (1995) provide empirical evidence that the correlation between dispersion of beliefs and turnover is positive and conclude that high trading volume is a good proxy for low consensus among informed traders.

specification check, in Table 3 Panel D we report the results where *turnover* is replaced with the standard deviation of analyst earnings forecasts divided by the absolute value of the mean forecast, where the latter is a commonly-used alternative proxy for dispersion in beliefs. As shown, we find a significant and positive coefficient on *SNetOptXStdForecasts*, consistent with the notion that dispersion in investor beliefs creates demand for soft information, ultimately resulting in a greater price responsiveness to soft information surprises for firms with higher dispersion in informed investor beliefs.

Table 4 Panel E presents the results where *numerical terms*, measured as the simple count of the number of numerical terms in the announcement divided by the number of words, are interactively included in the announcement period returns regression. For the Compustat sample, we find that *SNetOpt* interacted with numerical terms is positive and significant, consistent with the notion that providing more detailed, precise, and hard information (i.e., numbers) enhances the credibility of the net optimism concurrently expressed in the announcement, resulting in the net optimism being priced more.

Table 4 Panel F shows the differential market response to high-tech versus non-tech firms, where the high tech dummy variable is set equal to one for firms that fall into the Fama and French (1988) high tech industry portfolio definition.[31] The results suggest that there is a stronger market response to high tech firms' unexpected net optimism, although this is just barely significant for the Compustat sample. These findings are consistent with the notion that soft verbal data may be more important to conveying information to market participants in settings where the hard accounting data is less informative about the firms' economic realizations.[32] The results reported in Table 4 Panels G and H for, respectively, the forward PE-ratio, measured as the share price as of the current fiscal year end divided by expected earnings (IBES median forecast) for the following fiscal year, and historical R&D intensity, measured as R&D expense scaled by total expenditures, further support this notion of the market's responding more intensively to the soft information of higher growth

[31] Specifically, high tech firms include those with SIC codes of 3570-3579, 3622, 3660-3692, 3694-3699, 3810-3839, 7370-7372, 7373-7379, 7391, or 8730-8734.

[32] Lev and Zarowin (1999) argue that accounting data are less useful for firms engaged in innovative activities or that operate in more dynamic environments. One reason for this is that US accounting regulations require that firms expense all expenditures on R&D, leading to depressed accounting earnings even though the R&D expenditures may represent valuable investments in intangible assets that will generate growth in revenues and earnings into the future.

option, more intangibles-laden firms for which traditional historical accounting numbers are least informative.[33]

Table 4 Panel I pursues the notion of earnings informativeness further by examining the regression that interactively includes a measure of earnings quality, which we estimate using the data and methodology provided by Ecker, Francis, Kim, Olsson and Schipper (2006) (hereafter "EFKOS"). Higher values for the *EFKOS e-loading* factor represent lower levels of earnings quality. Lower earnings quality as captured by the EFKOS factor may result from factors that are innate to the firm's business model or from managerial attempts at obfuscation. Our results suggest that for lower earnings quality firms, the market response to net optimism is stronger, consistent with the notion that management provides soft information in order to compensate for hard accounting data that do not adequately capture the underlying economics of their firm's activities. Surprisingly, however, the market also weights earnings more heavily for firms that have higher EFKOS factors (i.e., lower quality earnings).

Table 4 Panel J presents the results from a multivariate regression that includes as independent variables all of the proxies that are not related to similar underlying constructs (and thus for which there is no strong theoretical correlation). Our findings support the incremental significance of each of the three constructs that we have separately examined: the extent of numerical terms used in the statement, which serves as a proxy for the verifiability of the information conveyed by management; turnover, which is highly correlated with analyst following and media coverage and that represents the revelation incentive effect of two-way communications; and earnings quality, a proxy for the demand for additional soft information when the alternative hard earnings news is not sufficient.

In untabulated results, we have also investigated the differential responsiveness of price to net optimism interacted with the market-to-book ratio as a proxy for "glamour" stocks, the level of competition in the industry as captured by a Herfindahl index, and certainty as a measure of ambiguity in management's choice of language.[34] None of these interacted variables are statistically significant.[35] Davis, et

---

[33] In an untabulated specification check, we find similar results when R&D intensity is measured as R&D expense scaled by sales.

[34] Benabou and Laroque (1992) suggest that managers have an incentive to strategically release manipulated soft information while claiming that the firm faces a highly uncertain environment. This

al. (2007) have previously documented that controlling for the simultaneous release of management earnings forecasts does not attenuate their results regarding the market's response to soft information, so we do not replicate their analyses here.

3.2 The Predictive Role of Net Optimism for Post-Announcement Drift

In this section, we examine whether the unexpected net optimism in management's earnings announcements is associated with the post-earnings announcement drift ("PEAD") phenomenon. In doing so, we extend the work of Davis, et al. (2007), who document only an announcement period market response to management-released soft information, and Engelberg (2007), who interprets his finding of a delayed price response to media-issued soft information to be indicative of higher information processing costs (i.e., relative to those associated with hard earnings news).

3.2.1 The Relation Between Net Optimism and Post-Announcement Returns

Following the hard earnings based PEAD literature (e.g., Ball and Brown (1968), Bernard and Thomas (1989)) and the media-expressed soft information results of Engelberg (2007), we hypothesize that there will be a partially delayed response to the soft information expressed in managerial earnings announcements, and thus that post-announcement abnormal returns will be associated with unexpected net optimism. For these tests, our dependent variable is defined as the 60-trading-day, size- and book-to-market-adjusted cumulative abnormal returns (CARs) for the period [+2, +62] relative to t=0 as the earnings announcement day.[36] The results for the baseline PEAD analyses are presented in Table 5A. As shown, net optimism is positive in all of the panels although its significance is attenuated in the IBES sample regressions when the surprise in net optimism is measured using the Diction algorithm.

---

way, if the manipulated information does not materialize as disclosed, managers can blame the unfulfilled expectations on the environmental uncertainty.

[35] Since our event window returns dependent variable is adjusted for the book-to-market effect, it's not entirely unexpected that the coefficient on the interaction of surprise net optimism and the market-to-book ratio is insignificant.

[36] Our results using GI are robust to alternative 70-, 90-, and 120-day post-announcement intervals.

We next consider whether it takes relatively longer for the market to incorporate soft information into asset prices than hard information.[37] Specifically, we estimate the following seemingly unrelated regression model,

$$\begin{aligned} \sum_{i=+2}^{62} AR_{jt+i} &= \beta_{10} + \beta_{1SUE} SUE_{jt} + \beta_{1Sent} SNetOpt_{jt} + \varepsilon_{1jt}, \\ \sum_{i=-1}^{62} AR_{jt+i} &= \beta_{20} + \beta_{2SUE} SUE_{jt} + \beta_{2Sent} SNetOpt_{jt} + \varepsilon_{2jt}, \end{aligned} \qquad (2)$$

and test the null hypothesis that $H_0 : \frac{\beta_{1Sent}}{\beta_{2Sent}} = \frac{\beta_{1SUE}}{\beta_{2SUE}}$. In other words, we test whether the stock price response during the post-announcement period as a fraction of the total stock price response during the combined announcement and post-announcement period is equivalent for soft and hard information, respectively. Table 5B shows that a larger fraction of the total response comes in the post-announcement period for the soft information variable than for the hard information variable, i.e. $\frac{\beta_{1Sent}}{\beta_{2Sent}} > \frac{\beta_{1SUE}}{\beta_{2SUE}}$ (except when using Diction sentiment in the Compustat sample) and the difference between the ratios is statistically significant in the IBES sample for both sentiment measures. These results are consistent with Engelberg (2007), who finds that soft information has greater predictability for returns at longer horizons. Our findings suggest that it is easier for the market to understand the pricing implications of SUE relative to soft information.

Table 5C examines the interaction of certainty with each of SUE and net optimism in the PEAD regressions, while also controlling for firm size. Consistent with the empirical findings of Vega (2006) and the "structural uncertainty theories" of PEAD (Brav and Heaton (2002)), we find that PEAD is greater for firms that face more uncertain environments. The results for this interacted variable are only statistically significant when net optimism is measured using the GI algorithm, however, consistent with other evidence that the GI net positivity wordlist is a better predictor of current and future stock returns than the wordlist associated with Diction's net optimism.[38] In short, we find that unexpected net optimism is

---

[37] We thank Jeremy Stein for suggesting this hypothesis test.

[38] In further untabulated analyses, we also consider the interactive role of turnover, analyst coverage, numerical terms, high tech industry participation, forward-looking P/E ratios, R&D intensity, and earnings quality ("EFKOS") on the predictive role of net optimism for post-announcement returns, however none of these interacted terms are significant.

significantly associated with PEAD, and this relationship varies with the uncertainty of the information environment but not across the firm characteristics examined.

3.2.2 Hedge Returns and the Net Optimism-PEAD Relation

Based upon the previous results suggesting a predictive role for net optimism in the post-announcement period, we document the hedge returns available from a net optimism-based trading strategy. Panel A of Table 6 presents benchmark hedge returns from going long (short) in firms in the highest (lowest) SUE terciles. Panel B of Table 6 presents the hedge results from going long (short) in firms that fall into both the highest (lowest) hard earnings surprise and highest (lowest) soft net optimism surprise terciles where the net optimism being used is the Diction-based measure. Both hedge strategies are implemented on a size (i.e., market capitalization) stratified basis, with large firms being defined as those in the 9^{th} and 10^{th} deciles, medium firms coming from deciles 6 through 8, and small firms consisting of the remaining firms from deciles 1 through 5. As shown in the furthest right-hand column of both panels, the hedge returns available from small- and medium-sized firm portfolios are statistically and economically significant for both the SUE and combined SUE-net optimism portfolio sorts (i.e., ranging from 1.7% to 3% for a 60-trading-day holding period, or approximately 6.8% to 12% annually). However, the returns available from the combined soft and hard earnings news strategy are considerably higher than those from the SUE only strategy for both the medium firm portfolio (8.4% versus 6.8% annualized) and especially for the small firm portfolio (16.4% versus 12% on an annualized basis). The untabulated returns available from the same strategy using the GI-measured unexpected net optimism are even higher than those from the Diction-based measure, ranging from 2.3% to 5.5% for a 60-trading-day holding period, or approximately 9.21% to 22% annually.

3.3 Summary

In this section we have revisited the *media*-based analyses of Tetlock, et al. (2008) and Engelberg (2007) using *management*-issued soft information and find that the market responds to both the hard earnings and soft information surprises in management earnings announcements. We have also extended the managerial earnings announcement mean short-window results of Davis, *et al.* (2007) by documenting the predictive role of management's soft information for post-announcement abnormal returns and by examining the conditions under which soft information is more informative to market participants. The latter results provide rare archival empirical support for cheap talk theoretical models.[39]

4. Soft Information and Idiosyncratic Volatility

In the preceding section, we have examined and extended the literature's findings with respect to the relation between soft information and the first moment of stock returns. In this section, we take the literature in a different direction by documenting the relation between unexpected net optimism and another linguistic characteristic, *certainty*, and abnormal idiosyncratic volatility. We examine the relation between these two dimensions of soft information and the second moment of stock returns after controlling for other previously documented determinants of idiosyncratic volatility.

We define *volatility* as the log of the sum of the squared abnormal returns over the event window of [-1, +1], where abnormal returns are calculated as above (i.e., as the firm's own daily return minus the contemporaneous return on a size- and book-to-market-matched portfolio) and we test for the association between our measures of soft information and volatility using the following specification:

$$\begin{aligned}&\log\left(\sum_{i=-1}^{1} AR_{jt+i}^2\right) = \gamma_0 + \gamma_1 \log\left(\sum_{i=2}^{4} AR_{jt-i}^2\right) + \gamma_2 \left|SUE_{jt}\right| + \gamma_3 \left|SUE_{jt}\right| \times I(SUE_{jt} < 0) \\ &+\gamma_4 \left|SNetOpt_{jt}\right| + \gamma_5 \left|SNetOpt_{jt}\right| \times I(SNetOpt_{jt} < 0) + \gamma_6 certainty_{jt} + \overline{\gamma}_5 \overline{\Upsilon}_{jt} + \varepsilon_{jt}.\end{aligned} \tag{3}$$

Given the well-known auto-regressive properties of idiosyncratic volatility, we control for the firm's own past volatility. In order to examine whether the soft information measures are capturing information that is incremental to the quantifiable riskiness of the firm's earnings and cash flows (i.e., valuation fundamentals), we also control for indicators of the volatility inherent in the firm's underlying business model. Specifically, we include a measure of past profitability, ROE, as well as the volatility of past profitability, Vol(ROE), both of which Pastor and Veronesi (2003) have found to be important drivers of idiosyncratic volatility. We follow Pastor and Veronesi (2003) and measure the Vol(ROE) for each firm as the residual variance

[39] Most of the empirical research related to cheap talk has been generated in experimental laboratory settings. One exception is Goetzmann, Ravid, Sverdlove and Pons-Sanz (2007) who examine the role of hard and soft information in the pricing of screenplays using an archival database of movie script sales.

from an AR(1) model of return on equity (ROE).[40] Because negative shocks may have a larger impact on the volatility of stock returns than positive shocks of the same absolute value,[41] we include a separate term capturing |SUE| interacted with an indicator variable for negative earnings surprises and a separate term capturing |SNetOpt| interacted with an indicator variable for negative net optimism surprises. The remaining control variables in the regression are described in Appendix B.

We first present the results from the volatility regression in Table 7A excluding ROE and Vol(ROE) since our sample is diminished significantly when we include these two variables. Of primary interest to our study is the result that the soft information *certainty* variable is significantly negatively associated with variance. The certainty measure is designed to capture management's tendency to speak in a resolute, complete, and straightforward manner. By implication, lower levels of certainty are associated with less resolute, and thus more obtuse or obfuscated use of, language. Our findings of a negative association between certainty and unexpected idiosyncratic volatility suggest that management's use of more wavering language in their earnings press release leads to greater uncertainty regarding the level or riskiness of the firm's future cash flows, resulting in higher volatility in the firm's share price during the 3-day announcement period window. Our certainty measure is negatively correlated with Vol(ROE) and the dispersion of analyst forecasts (i.e., more certain language is associated with lower variance profitability and lower dispersion in informed analysts' beliefs), suggesting that the linguistic measure captures some otherwise quantifiable aspects of the riskiness in earnings and cash flows. However, the finding that the linguistic measure has incremental explanatory power for volatility suggests that this variable also captures other aspects of the uncertainty in the firm's environment, such as changes in circumstances that are not reflected in past hard information realizations.[42]

The surprise in net optimism is also a significant determinant of announcement interval volatility, however this result only holds for the Compustat

[40] We refer the reader to Pastor and Veronesi (2003) for a more detailed definition of the variables.
[41] This phenomenon is most often interpreted as the leverage effect unveiled by Black (1976). Several GARCH volatility models allow for this effect, including the EGARCH model of Nelson (1991), and the GJR model of Glosten, Jagannathan and Runkle (1993), amongst others.
[42] For example, the low certainty score of the example provided in Panel D of Appendix A is partly driven by a recent bankruptcy declaration of one of the firm's distributors. Other low certainty scores were recorded by biotech firms who face uncertainty regarding the outcomes of clinical trials and in the pharmaceutical approval process.

sample. With respect to the control variables, past volatility is the most important determinant of announcement period volatility, as expected, and consistent with prior studies we find that volatility is an increasing function of the absolute value of SUE. Firms with lower quality earnings (i.e., higher EFKOS e-loading factors), higher levels of media exposure, higher levels of analyst following, and that voluntarily provide more financial statements all have higher levels of abnormal volatility during the announcement window, while large firms, firms with longer press releases, and firms issuing announcements in the post-RegFD period (only for the Compustat sample) exhibit lower levels of volatility. Also consistent with prior results (e.g., Pastor and Veronesi (2003)), we document that firms with high MB ratios have higher idiosyncratic volatility, while firms that pay dividends or that are highly leveraged have lower volatility. Firm age is not a significant determinant of idiosyncratic volatility for either of our Compustat or IBES samples.

In Table 7B we present the results that include the ROE and Vol(ROE) variables that Pastor and Veronesi (2003) found to be important determinants of idiosyncratic volatility. Consistent with their results, we find that the volatility of ROE is positively correlated with idiosyncratic volatility and the level of profitability is negatively related to the firm's volatility, while the sign and significance of the coefficients on our other control variables generally remain unchanged. Overall our results suggest that, after controlling for fundamental indicators of the firm's inherent economic riskiness, market volatility is incrementally responsive to soft information measures of uncertainty derived from the text of managers' earnings announcements.

5.1 Certainty as a Predictor of Post-Announcement Idiosyncratic Volatility

In Table 8 we report the results for equation (3) estimated using the 60-trading-day post-announcement period [+2, +62] stock volatility and using past volatility measured over [-62,-2] as the corresponding control variable. Similar to the announcement period volatility results, *certainty* takes a negative coefficient and is a significant predictor of volatility for both samples. In addition, unexpected net optimism is significantly associated with post-announcement idiosyncratic volatility for both samples. Thus, not only are the soft information variables significantly related to contemporaneous announcement period idiosyncratic volatility, they are also leading indicators for post-announcement volatility. Most of the results for the control variables in the longer window are similar to those in the announcement period, except that the time trend variable is negative and significant (consistent with declining idiosyncratic volatility during the period covered by our data), and the *REG_FD* variable and the length of the announcement are no longer significant. Also similar to the announcement period results, the soft information variables all remain significant for both samples even after controlling for the Pastor and Veronesi (2003) fundamental indicators of economic volatility as reported in Table 8B.

Overall, our results show that the certainty of manager's language is inversely related to the level of idiosyncratic volatility both in the short window announcement period and in the intermediate post-announcement drift interval. Because our findings are robust to the inclusion of proxies for the firm's past realized levels of fundamental economic uncertainty, our results may contribute to the further development of cheap talk theories which have heretofore largely neglected the second moment of soft information.[43] We show that a linguistic measure of certainty captures some incremental aspect of managerial wavering or indirect use of language that in turn conveys information regarding the uncertainties facing the firm, and that this soft information increases investor uncertainty about the stock's valuation.

5. Summary and Conclusion

Prior research has established a general link between soft information, captured with the linguistic measures of negativity or pessimism, and stock returns. Most of this prior literature examines *media*-released information and the association between soft information and the first moment of stock returns. Our study compliments and extends the existing studies by examining the conditions under which *management*-issued soft information is incorporated into prices, both in the short-window announcement period and in the intermediate term, post-announcement interval. Consistent with cheap talk theories, we find that under conditions that induce the provision of informative soft information rather than mere noise (e.g., where stock turnover and analyst scrutiny are high), the market responds more to the surprise net optimism in managerial announcements. We also find that in circumstances where the hard earnings information is less informative (e.g., for high

[43] One exception is Benabou and Laroque (1992) who show that if managers observe a noisy signal of the truth, i.e. qualify their sentiment with high uncertainty, managers will convey strategically distorted messages. However, our empirical evidence does not support this theory, since our certainty measure does not affect the significance of sentiment in explaining contemporaneous abnormal returns.

tech firms with more complex business models and for firms with lower earnings quality as captured by the EFKOS e-loading factor), the market responds more to the complimentary soft information. Consistent with the view that soft information is more difficult for market participants to process, we find that it takes longer for soft information than for hard information to be incorporated into prices. We extend the soft information literature further by examining the association between another linguistic measure, certainty, and the second moment of returns, abnormal idiosyncratic volatility. We find that the level of certainty expressed in management's earnings announcement is inversely related to idiosyncratic volatility during the announcement window and that this linguistic characteristic is also a leading indicator for post-announcement 60-trading-day abnormal volatility. Thus, the use of less resolute language in management's earnings announcements is incrementally associated with stock valuation uncertainty, and these results hold even after controlling for fundamental indicators of the past uncertainty in the firm's economic environment. Taken together, our findings suggest that management-conveyed soft information plays an important role in the price discovery process that is incrementally informative and complementary to the simultaneously released hard earnings news that has been the subject of decades of prior research.

Appendix A
Excerpts from Earnings Announcement Texts

The following panels present excerpts from firms' earnings announcements exhibiting the linguistic characteristics being exemplified (i.e., high optimism, high pessimism, high and low uncertainty, respectively). We highlight in the text the Diction dictionary words associated with the underlying linguistic construct.

Panel A – High Optimism: Home Depot, November 15, 2005, netopt = 3.7928-0.438=3.3548

ATLANTA, Nov. 15 /PRNewswire-FirstCall/ -- The Home Depot(R), the world's largest home improvement retailer, today reported third quarter fiscal 2005 net earnings of $1.5 billion, $0.72 per diluted share, up 20 percent, compared with $1.3 billion, $0.60 per diluted share, for the third quarter of fiscal 2004. Sales for the period increased $2.0 billion, or 10.5 percent, to $20.7 billion. Growth in comparable store sales was 3.6 percent. 'The execution of our strategy and the focus on our fundamentals has enabled us to consistently and predictably deliver strong results,' said Bob Nardelli, chairman, president & CEO. 'We have stayed on strategy, effectively managed our business and produced solid earnings growth through the hard work and dedication of our 325,000 associates.'

'We continued to drive productivity throughout our business, and are well on our way to becoming the low cost provider in our industry. During the quarter we continued to use our strong financial condition to invest in the business and return cash to our shareholders,' said Carol Tome, executive vice president and CFO. At the end of the third quarter, the company reported total assets of $44.7 billion, total stockholders' equity of $25.8 billion and return on invested capital of 21.8 percent. In the third quarter, the company repurchased $868 million or 21.8 million shares under its share repurchase program. Since the inception of its share repurchase program, the company has repurchased $9.5 billion or 272 million shares under its $11 billion authorization. The company has repurchased about 12 percent of its outstanding shares since 2002. The company lifted its fiscal 2005 sales growth guidance from 9-12 percent to 10-12 percent and increased its earnings per share growth guidance from 14- 17 percent to 17-18 percent.

Enhancing the Core

By broadening its assortment and adding new, innovative and distinctive merchandise that provides tremendous customer value, The Home Depot achieved a record average ticket of $58.92, representing an increase of 6.1 percent compared to the third quarter of last year.

'During the quarter our merchants did a great job of adding innovative and distinctive products to our stores. Our strategy of enhancing the core through distinction and innovation is working as evidenced by the highest average ticket in our company's history. We saw average ticket growth across the store with real strength in kitchen and bath,' said Tom Taylor, executive vice president, Merchandising and Marketing. 'The active hurricane season during the quarter showcased our merchandising and operational flexibility. We took extraordinary steps to take care of our communities, customers and associates and we were the first retailer to deliver emergency-type merchandise to the affected areas. By the end of the quarter, we directed over 4,000 truckloads of merchandise to the Gulf Coast and Florida regions,' added Taylor.

Panel B – High Pessimism: National Steel, January 24, 2001, netopt = 0.4724-2.296= -1.8236

National Steel Corporation (NYSE: NS) today reported a net loss of $ 83.4 million, or $ 2.02 per diluted common share for the fourth quarter of 2000. This compares to a restated net loss of $ 3.2 million, or $ 0.08 per diluted common share for the fourth quarter of 1999. The Company is restating its financial statements for 1998, 1999 and the first three quarters of 2000, as discussed below. Net sales for the quarter amounted to $ 651.6 million on steel shipments of 1,434,000 tons which compares to net sales of $ 785.7 million on steel shipments of 1,686,000 tons in the year earlier quarter. Shipments and revenues were negatively impacted in the fourth quarter by continued high levels of steel imports combined with a general slowdown in the U.S. economy which has affected the demand for our products. We have also seen a dramatic decrease in the spot market selling price for steel products which has further impacted our financial performance. Our production costs have been negatively impacted by higher energy costs, especially natural gas, and reduced levels of production at our steel making facilities given the weakness in demand for our products.

For the full year 2000, the Company reported a loss of $ 129.8 million, or $ 3.14 per diluted common share as compared to a restated net loss of $ 28.6 million, or $ 0.69 per diluted common share for the year 1999. Net sales for the year 2000 rose by $ 25.5 million to $ 2,978.9 million and steel shipments increased 2.4% to 6,254,000 tons.

"We are very disappointed by our performance and financial results for 2000, particularly in the second half of the year," said Yutaka Tanaka, chairman and chief executive officer. "Several factors including high import levels and a weakening U.S. economy have severely impacted the steel industry. We continue to take the necessary steps to reduce our costs and ensure adequate liquidity during this very severe downturn," he said.

FINANCIAL POSITION AND LIQUIDITY

Total liquidity from cash and available short-term borrowing facilities amounted to $ 118 million at December 31, 2000 as compared to $ 253 million at September 30, 2000. The primary reasons for the decline in liquidity during the fourth quarter 2000 were the continued net losses, scheduled debt repayments and a reduction in availability under the accounts receivable securitization credit facility.

…

OUTLOOK

The Company's outlook for the near-term remains pessimistic. Forecasted shipments for the first quarter 2001 are expected to be slightly lower than the fourth quarter 2000, impacted by the automotive market which has weakened further and the construction market which has been impacted by weather conditions. Steel imports and the effect of a slowing U.S. economy continue to negatively impact spot market pricing for our products. The Company believes that it has adequate liquidity for the near-term and is in continuing discussions with its lead banks to ensure future compliance with all financial covenants. Capital spending will be dramatically reduced during the first quarter 2001 to approximately $ 14 million and additional measures are being taken to control our cash outflows.

Panel C – High Certainty: Raytheon, July 28, 2005, certainty = 43.5456

WALTHAM, Mass., July 28, 2005 /PRNewswire-FirstCall/ -- Raytheon Company (NYSE:RTN) reported second quarter 2005 income from continuing operations of $233 million or $0.51 per diluted share compared to a loss from continuing operations of $94 million or $0.22 per diluted share in the second quarter 2004. Second quarter 2004 income from continuing operations, excluding the effect of charges for the settlement of a class action shareholder lawsuit and the early retirement of debt, was $152 million or $0.35 per diluted share. Second quarter 2005 income from continuing operations was higher due to better operating results in the Government and Defense businesses and at Raytheon Aircraft Company (RAC) combined with lower interest expense.

'I continue to be pleased with performance throughout the Company,' said William H. Swanson, Raytheon's Chairman and CEO. 'The strength of our bookings and record backlog demonstrate that the Company is well positioned for future growth.'

Second quarter 2005 net income was $201 million or $0.44 per diluted share compared to a net loss of $108 million or $0.25 per diluted share in 2004. Net income for the second quarter of 2005 included a $32 million after-tax loss in discontinued operations or $0.07 per diluted share, primarily attributable to foreign tax related matters, versus a $14 million after-tax loss or $0.03 per diluted share in 2004.

Net sales for the second quarter 2005 were $5.4 billion, up 10 percent from $4.9 billion in the comparable period in 2004. Government and Defense sales for the quarter (after the elimination of intercompany sales) increased 8 percent to $4.5 billion from $4.2 billion in the comparable quarter. RAC sales for the quarter increased 21 percent to $687 million from $570 million in the 2004 comparable quarter.

Free cash flow from continuing operations for the second quarter 2005 was $736 million versus $820 million for the comparable period in 2004, a decrease primarily due to timing of collections. Year-to-date free cash flow was $398 million versus $620 million for the comparable period in 2004, a decrease primarily due to a $200 million discretionary cash contribution to the Company's pension plans made in the first quarter of 2005. Free cash flow is defined by the Company as operating cash flow less capital spending and internal use software spending.

During the second quarter of 2005, the Company repurchased 3.6 million shares of common stock for $139 million as part of the Company's previously announced $700 million share repurchase program, bringing the total shares of common stock repurchased year-to-date to 4.9 million for $192 million.

Net debt was $4.6 billion at the end of the second quarter 2005 and at the end of 2004.

Panel D – Low Certainty: Play-By-Play Toys & Novelties, June 14, 1996, certainty = 9.1072

Play-By-Play Toys & Novelties, Inc. (Nasdaq: PBYP) today announced results for the 1999 fiscal third quarter and nine months ended April 30, 1999. Net sales for the third quarter of fiscal1999 decreased 4.2% to $35.7 million, from $37.3 million reported for the third quarter of fiscal 1998. Net loss for the third quarter of 1999 was $3.6 million, or $0.49 per share, compared with net income of $1.7 million, or $0.22 per diluted share, for the comparable period last year. This loss includes the write-off of $3.3 million of accounts receivable following the commencement of insolvency proceedings by the Company's Mexico distributor and approximately $362,000 of severance pay related to restructuring initiatives and personnel reductions, all of which are included in selling, general and administrative expenses. Excluding the one-time charge for bad debt expense and severance payments, the Company's net loss would have been $1.2 million, or $0.17 per share, for the third quarter of 1999. The principal factors contributing to the overall decline in sales and earnings for the third quarter was a decrease in domestic retail sales and international amusement sales of approximately $1.7 million and $1.0 million, respectively, offset by an increase in international retail sales of $1.6 million over the comparable period a year ago. In addition to the previously mentioned items, the Company's earnings were impacted by weaker gross margins resulting from increased licensing costs as a percentage of sales in domestic retail and in Europe, and sales of closeout merchandise at reduced margins in retail and amusement early in the quarter. Net sales for the nine months ended April 30, 1999 decreased 5.3% to $119.8 million, from $126.6 million, reported for the comparable period last year. Net loss for the first nine months of 1999 including the write-off was $4.1 million, or $0.56 per share, compared with net income of $5.2 million, or $0.77 per diluted share, for the comparable period last year. Raymond Braun, President and Chief Operating Officer of Play-By-Play, commented, "We made significant progress against several key strategic initiatives and we remain aggressively focused on further reducing operating costs and overhead expenses in an effort to get the Company's overall structure in-line with the current business environment. Our amusement business continues to be strong both internationally and domestically. The economic weakness and uncertain business environment in Latin America continues to negatively impact the Company's results and efforts relative to that market. Despite this, we are now seeing early signs of economic recovery and our long-term outlook for Latin America remains positive. Additionally, our sales approach to Latin America has changed. With the purchase of Caribe Marketing in Puerto Rico, we have positioned the Company to sell direct in several countries allowing us to increase sales volumes within those countries while significantly reducing our concentration of credit risk. We are further reducing credit risk through more conservative sales terms and we have forged relationships with financially strong distributors in certain key countries. As a result of these changes, we expect more moderate growth in Latin America over the next two years; however, our credit risk will be substantially lower, which should translate into a greater and more stable contribution to earnings in the long-term." Mr. Braun further commented, "The domestic retail market for traditional toys continues to experience weakness and we have made significant strategic changes in this market as well. We continue to focus on products similar to those in our other distribution channels including plush, novelties and interactive toys, with a better mix of price points, and with fewer promotional items. Our strategy for the remainder of calendar 1999 includes the elimination of television advertising, which should reduce selling costs by approximately $2.5 million. While this will result in lower retail sales in 1999 compared to 1998, the contribution to earnings from retail should improve. Based on the implementation of the strategy discussed above, we believe calendar 2000 will reflect improvement in both sales and profitability." Arturo Torres, Chairman of the Board and Chief Executive Officer of Play-By-Play commented, "We undertook a number of important changes this quarter, including an executive and management restructuring, personnel reductions and other strategic improvements aimed at keeping Play-By-Play successful within the highly competitive and dynamic toy industry. We are encouraged by the initial impact of these changes but recognize that additional time and effort will be needed to complete the process."

Appendix B
Summary of Variable Definitions

Variable	Definition
Soft Information Variables	
Optimism	Percentage number of words in a firm's quarterly earnings announcement that are optimism-increasing (i.e., contained in the praise, satisfaction, or inspiration dictionaries)
Pessimism	Percentage number of words in a firm's quarterly earnings announcement that are optimism-decreasing (i.e., contained in the blame, hardship, or denial dictionary definitions)
Net Optimism	Optimism minus pessimism
Certainty	Certainty, is a normalized variable that indicates the degree of "resoluteness", "inflexibility", and "completeness" in the firm's quarterly earnings announcement. We redefine the Diction 2.0 definition of certainty to be [Tenacity + Leveling + Collectives + Insistence + Numerical Terms] - [Ambivalence + Self Reference + Variety]
Other Variables	
SUE	Earnings surprise = $\frac{actual - forecast}{std(actual - forecast)}$
CARs	Size- and book-to-market-adjusted cumulative abnormal returns defined alternatively over the earnings announcement window [t-1,t+1] or the post-announcement drift period [t+2,t+62] relative to the *t=0* earnings announcement day.
Volatility	We measure the volatility of abnormal returns during the event window as the logarithm of the sum of squared abnormal returns during the [t-1, t+1] and [t+2, t+62] event windows.
Time Trend	= 1 for 1st calendar quarter of 1998, increased by 1 for each calendar quarter thereafter
RegFD	Indicator=1 for firm quarters that occur after October 23, 2000
Financial Statements	We create a count variable, *Financial Statements*, which is incremented by one for each voluntary disclosure of the following items within the earnings announcements: a cash flow statement, an income statement, a balance sheet, and a tabulated summary of financial highlights
Total Words	Natural log of the total number of words contained in the earnings announcement
Log(Analyst+1)	Analyst is computed using IBES data and it is equal to the number of analysts that post an earnings estimate for the current quarter
Forecast Dispersion	We use IBES to estimate this variable, and define it as the standard deviation of forecasts across analysts divided by the absolute value of the median forecast. We require firms to at least have two forecast estimates.
Turnover	The average of the natural log of de-trended turnover (the daily volume of shares traded divided by stock outstanding) cumulated over the pre-announcement period [t-62, t-2]. In order to present a pooled regression of NYSE/AMEX and Nasdaq firms, we follow the common heuristic of dividing the Nasdaq firms' volume by two (Atkins and Dyl (1997) and Dyl and Anderson (2005)). We de-trend turnover using Campbell, et al. (1993) method of calculating the turnover's trend as the rolling average of the past 60 trading days. We add back the mean of turnover to our de-trended

	measure, so that the units are economically meaningful.
Recent Media Coverage	The number of times a firm is mentioned in the headline or lead paragraph of an article from newswire services in the previous 60 trading days before the earnings announcement date [t-62,t-2]. We only take into account publications that have over 500,000 current subscribers using Factiva. The list of data sources is: The Wall Street Journal (all editions), Associated Press Newswire, the Chicago Tribune, the Globe and Mail, Gannett News Service, the Los Angeles Times, the New York Times, the Washington Post, USA Today and all Dow Jones newswires.
High Tech	Indicator = 1 if dnum 3570-3579, 3622, 3660-3692, 3694-3699, 3810-3839, 7370-7372, 7373-7379, 7391, 8730-8734
PE Ratio	We use IBES data to construct this ratio on an annual basis. We take the average price of the firm during the fiscal year divided by the expected earnings for that year.
R&D Expenses	We estimate the annual R&D expenses (data 4 in the quarterly Compustat tape) as a fraction of total expenditures (data 1 plus data 4 in the quarterly Compustat tape).
EFKOS e-Loading	Is obtained by regressing the daily excess return of firm *i* on EFKOS factor as well as the Fama-French three factors (SML, HML, Market Return). We allow the loading to change over time and estimate the coefficient using all non-earnings announcement days in the previous 365 calendar days (only for stocks with at least 100 data points in that period) before the earnings announcement.

References

Atkins, Allen, and Edward A. Dyl, 1997, Market Structure and Reported Trading Volume: Nasdaq versus the NYSE, *Journal of Financial Research* 20, 291-304.

Ball, Ray, and Phillip Brown, 1968, An Empirical Evaluation of Accounting Income Numbers, *Journal of Accounting Research* 6, 159-78.

Barber, Brad M., and John D. Lyon, 1997, Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics, *Journal of Financial Economics* 43, 341-372.

Benabou, R., and G. Laroque, 1992, Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility, *Quarterly Journal of Economics* 107.

Bernard, Victor, and Jacob Thomas, 1989, Post-Earnings-Announcement Drift, *Journal of Accounting Research* 27, 1-36.

Bhattacharya, Nilabhra, Ervin Black, Theodore Christensen, and Chad Larson, 2003, Assessing the Relative Informativeness and Permanence of Pro Forma Earnings and GAAP Operating Earnings, *Journal of Accounting and Economics* 36, 285-319.

Black, Fischer, 1976, Studies of Stock Price Volatility Changes, Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section.

Bligh, Michelle, and Gregory Hess, 2007, The Power of Leading Subtly: Alan Greenspan, Rhetorical Leadership, and Monetary Policy, *The Leadership Quarterly* 18, 87-104.

Bradshaw, Mark, and Richard Sloan, 2002, GAAP versus The Street: An Empirical Assessment of Two Alternative Definitions of Earnings, *Journal of Accounting Research* 40, 41-66.

Brav, Alon, and J. B. Heaton, 2002, Competing Theories of Financial Anomalies, *Review of Financial Studies* 15, 575-606.

Burgstahler, David C., and Ilia D. Dichev, 1997, Earnings Management to Avoid Earnings Decreases and Losses, *Journal of Accounting and Economics* 24, 99-126.

Campbell, Jon , Sanford Grossman, and Jiang Wang, 1993, Trading Volume and Serial Correlation in Stock Returns, *Quarterly Journal of Economics* 108, 905-939.

Cenesizoglu, T., and A. Timmermann, 2008, Is the Distribution of Stock Returns Predictable?, *unpublished working paper, UCSD and HEC Montreal*.

Chan, Wesley S., 2003, Stock Price Reaction to News and No-News: Drift and Reversal After Headlines, *Journal of Financial Economics* 70.

Crawford, Vincent, and Joel Sobel, 1982, Strategic Information Transmission, *Econometrica* 50, 1431-1451.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the Characteristics of Cross-Sectional Variation in Common Stock Returns, *Journal of Finance* 52, 1-33.

Das, Sanjiv, and Mike Chen, 2007, Yahoo! for Amazon: Sentiment Extraction From Small Talk on the Web, *Management Science* 53, 1375-1388.

Das, Sanjiv, Asis Martinez-Jerez, and Peter Tufano, 2005, e-Information: A Clinical Study of Investor Discussion and Sentiment, *Financial Management* 34, 103-137.

Davis, Angela, Jeremy Piger, and Lisa Sedor, 2007, Beyond the Numbers: Managers' Use of Optimistic and Pessimistic Tone in Earnings Press Releases, *unpublished working paper*.

Degeorge, Francois, Jayendu Patel, and Richard Zeckhauser, 1999, Earnings Management to Exceed Thresholds, *Journal of Business* 72, 1-33.

Diamond, Douglas, and Robert Verrecchia, 1987, Constraints on Short-Selling and Asset Price Adjustment to Private Information, *Journal of Financial Economics* 18, 277-311.

Digitext Inc., 2000, Diction 5.0: The Text-Analysis Program, (Austin, Texas).

Dyl, Edward A., and Anne-Marie Anderson, 2005, Market Structure and Trading Volume, *Journal of Financial Research* 28, 115-131.

Easton, H., 1940. *Word Frequency Dictionary* (Dover, New York).

Ecker, Frank, Jennifer Francis, Irene Kim, Per Olsson, and Katherine Schipper, 2006, A Returns-Based Representation of Earnings Quality, *The Accounting Review* 81, 749-780.

Ehrmann, Michael, and Marcel Fratzscher, 2007, Communication by Central Bank Committee Members: Different Strategies, Same Effectiveness, *Journal of Money, Credit & Banking* 39, 509-541.

Engelberg, Joseph, 2007, Costly Information Processing: Evidence from Earnings Announcements, *unpublished working paper, Northwestern University*.

Fama, Eugene, and Kenneth French, 1988, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246-273.

Fama, Eugene, and Kenneth French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427-465.

Foster, F. Douglas, and S. Viswanathan, 1996, Strategic Trading When Agents Forecast the Forecasts of Others, *Journal of Finance* 51, 1437-1478.

Glosten, Lawrence, Ravi Jagannathan, and David Runkle, 1993, On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *Journal of Finance* 48, 1779-1801.

Goetzmann, William N., S. Abraham Ravid, Ronald Sverdlove, and Vicente Pons-Sanz, 2007, The Role of Soft and Hard Information in the Pricing of Assets and Contract Design: Evidence from Screenplay Sales, *unpublished working paper, Yale School of Management*.

Graham, John, Campbell Harvey, and Shiva Rajgopal, 2006, Value Destruction and Financial Reporting Decisions, *Financial Analysts Journal* 62, 27-39.

Grossman, Sanford J., 1981, The Role of Warranties and Private Disclosure About Product Quality, *Journal of Law and Economics* 24, 461-483.

Grossman, Sanford J., and J. E. Stiglitz, 1980, On the Impossibility of Informationally Efficient Markets, *American Economic Review* 70, 393-408.

Harris, Milton, and Arturo Raviv, 1993, Differences of Opinion Make a Horse Race, *Review of Financial Studies* 6, 473-506.

Hong, Harrison, and Jeremy Stein, 2003, Differences of Opinion, Short-Sales Constraints, and Market Crashes, *Review of Financial Studies* 16, 487-525.

Kandel, Eugene, and Neil Pearson, 1995, Differential Interpretation of Public Signals and Trade in Speculative Markets, *Journal of Political Economy* 103, 831-872.

Krishna, Vijay, and John Morgan, 2004, The Art of Conversation: Eliciting Information from Experts Through Multi-Stage Communication, *Journal of Economic Theory* 117, 147-179.

Krishna, Vijay, and John Morgan, 2008, Cheap Talk, in Steven N. Durlauf, and Lawrence E. Blume, eds.: *The New Palgrave Dictionary of Economics* (Palgrave Macmillan).
Lev, Baruch, and Zarowin, 1999, The Boundaries of Financial Reporting and How to Extend Them, *Journal of Accounting Research* 37, 353-385.
Li, Feng, 2006, Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports?, *unpublished working paper, University of Michigan*.
Ljungqvist, Alexander, Christopher Malloy, and Felicia Marston, 2007, Rewriting History, *working paper, NYU*.
Lucca, D. O., and F. Trebbi, 2008, Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements, *unpublished working paper*.
Milgrom, Paul, 1981, Good News and Bad News: Representation Theorems and Applications, *Bell Journal of Economics* 12, 380-391.
Nelson, Daniel B., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.
Ober, S., J. Zhao, R. Davis, and M. Alexander, 1999, Telling It Like It Is: The Use of Certainty In Public Business Discourse, *The Journal of Business Communication* 36, 280-300.
Pastor, Lubos, and Pietro Veronesi, 2003, Stock Valuation and Learning About Profitability, *Journal of Finance* 58, 1749-1789.
Peress, Joel, 2008, Media coverage and Investors' Attention to Earnings Announcements, *unpublished working paper*.
Petersen, Mitchell, 2004, Information: Hard and Soft, *unpublished working paper, Northwestern University*.
Petersen, Mitchell, 2008, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies* forthcoming.
Schrand, Catherine M, and Beverly Walther, 2000, Strategic Benchmarks in Earnings Announcements: The Selective Disclosure of Prior-Period Earnings Components, *The Accounting Review* 75, 151-177.
Sobel, Joel, 1985, A Theory of Credibility, *Review of Economic Studies* 52, 557-573.
Stein, Jeremy C., 2002, Information Production and Capital Allocation: Decentralized vs. Hierarchical Firms, *Journal of Finance*.
Stocken, Phillip, 2000, Credibility of Voluntary Disclosure, *RAND Journal of Economics* 31, 359-374.
Stone, Philip, Dexter Dunphy, Marshall Smith, and Daniel Ogilvie, 1966. *General Inquirer: A Computer Approach to Content Analysis* (MIT Press, Cambridge).
Tetlock, Paul C., 2007, Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *Journal of Finance* 62, 1139-1168.
Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More Than Words: Quantifying Language to Measure Firms' Fundamentals, *Journal of Finance* 63, 1437-1467.
Vega, Clara, 2006, Stock Price Reaction to Public and Private Information, *Journal of Financial Economics* 82, 103-133.
Veldkamp, Laura, 2006, Media Frenzies in Markets for Financial Information, *American Economic Review* 96, 577-601.
Wang, F. Albert, 1998, Strategic Trading, Asymmetric Information, and Heterogeneous Prior Beliefs, *Journal of Financial Markets* 1, 321-352.

Yuthas, K., R. Rogers, and J. Dillard, 2002, Communicative Action and Corporate Annual Reports, *Journal of Business Ethics* 41, 141-157.

Table 1A: Soft Information Sample Statistics

In this table we present summary statistics for the following variables estimated using earnings press releases and Diction 5.0 textual-analysis program: Optimism, the average number of words per one hundred words in a firm's quarterly earnings announcement that are optimism-increasing; pessimism, the average number of words per one hundred words in a firm's quarterly earnings announcement that are optimism-decreasing; Netopt, optimism minus pessimism; ΔNetOpt, change in Netopt from this quarter to the previous quarter; and certainty, is a normalized variable that indicates the degree of "resoluteness", "inflexibility", and "completeness" in the firm's quarterly earnings announcement. The summary statistics are calculated using 3,764 (2,610) firms sampled during earnings announcement dates from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. For a detailed description of the variables please refer to the Appendix.

	Mean	Std. Dev.	25 Percentile	75 Percentile
		Compustat Sample		
Optimism	1.3033	0.9790	0.6250	1.7654
Pessimism	0.6431	0.6302	0.2000	0.9000
NetOpt	0.6602	1.2199	0.0000	1.2832
ΔNetOpt	-0.0194	1.2716	-0.6318	0.5888
Certainty	31.1003	10.1311	23.8380	37.2827
		IBES Sample		
Optimism	1.3151	0.9677	0.6444	1.7708
Pessimism	0.6064	0.5861	0.2000	0.8494
NetOpt	0.7089	1.1851	0.0000	1.3158
ΔNetOpt	-0.0237	1.2359	-0.6098	0.5664
Certainty	31.2052	9.8576	24.1297	37.2110

Table 1B: Soft Information Correlation Matrix

We estimate the average quarterly bi-variate correlations of the variables used in the empirical tests. The bivariate correlations are calculated using 3,764 (2,610) firms sampled during earnings announcement dates from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. Optimism is the average number of words per one hundred words in a firm's quarterly earnings announcement that are optimism-increasing; pessimism is the average number of words per one hundred words in a firm's quarterly earnings announcement that are optimism-decreasing; Netopt is optimism minus pessimism; ΔNetOpt is the change in Netopt from this quarter to the previous quarter; and certainty, is a normalized variable that indicates the degree of "resoluteness", "inflexibility", and "completeness" in the firm's quarterly earnings announcement. For a detailed description of the variables please refer to the Appendix.

	Optimism	Pessimism	NetOpt	ΔNetOpt	Certainty	SUE
			Compustat Sample			
Optimism	1					
Pessimism	-0.1075	1				
NetOpt	0.8581	-0.6029	1			
ΔNetOpt	0.4541	-0.2843	0.5113	1		
Certainty	-0.1241	-0.0811	-0.0577	-0.0292	1	
SUE	0.0406	-0.0822	0.075	0.0562	0.0312	1
			IBES Sample			
Optimism	1					
Pessimism	-0.1097	1				
NetOpt	0.8708	-0.5842	1			
ΔNetOpt	0.4407	-0.2756	0.4962	1		
Certainty	-0.1015	-0.0867	-0.04	-0.0294	1	
SUE	0.0639	-0.1075	0.1053	0.0582	0.0143	1

Table 2A: Descriptive Statistics – Compustat Sample

In this table we present summary statistics for the variables used in the empirical tests. The variables are defined in Appendix B and the summary statistics are calculated using 3,764 firms sampled during earnings announcement dates from January, 1998 to July, 2006, and in total there are 21,580 firm-quarter observations.

	Mean	Std. Dev.	25 Percentile	75 Percentile
3-Day CAR	0.0031	0.0875	-0.0312	0.0352
Post announcement 60-Day CAR	0.0008	0.1865	-0.0788	0.0765
SUE	0.0559	1.0005	-0.3147	0.4577
SUE_IBES	0.2455	1.1522	-0.1618	0.7201
NEGV	0.2310	0.4215	0.0000	0.0000
Log(Market Capitalization)	19.8413	1.9443	18.4475	21.1119
Analyst following	3.7087	5.0108	0.0000	6.0000
Log(1+analyst)	1.0612	0.9808	0.0000	1.9459
Turnover	0.6292	0.6715	0.2013	0.7987
Media Coverage	20.5148	23.0481	3.0000	43.0000
EFKOS e-Loading factor	0.0970	0.5024	-0.1951	0.2877
Hightech indicator variable	0.1627	0.3691	0.0000	0.0000
Numerical Terms	81.3385	30.2652	58.9110	100.4960
Financial Statements	1.7582	0.9497	1.0000	2.0000
Total Words	824.7232	556.2340	440.0000	1064.0000
Log(Total Words)	6.5124	0.6503	6.0868	6.9698
R&D Expenditures	0.0000	0.2940	0.0000	0.3523

Table 2B: Descriptive Statistics – IBES Sample

In this table we present summary statistics for the variables used in the empirical tests. The variables are defined in Appendix B and the summary statistics are calculated using 2,610 firms sampled during earnings announcement dates from January, 1998 to July, 2006, and in total there are 13,907 firm-quarter observations.

	Mean	Std. Dev.	25 Percentile	75 Percentile
3-Day CAR	0.0020	0.0813	-0.0316	0.0360
Post announcement 60-Day CAR	0.0027	0.1615	-0.0695	0.0750
SUE	0.0666	0.9993	-0.3084	0.4742
SUE_IBES	0.2459	1.1507	-0.1625	0.7212
NEGV	0.2063	0.4047	0.0000	0.0000
Log(Market Capitalization)	20.4897	1.6846	19.2942	21.5673
Analyst following	5.8174	5.1841	2.0000	8.0000
Log(1+analyst)	1.6694	0.7009	1.0986	2.1972
Turnover	0.7380	0.6798	0.2941	0.9410
Media Coverage	22.6295	23.0688	4.0000	57.0000
EFKOS e-Loading factor	0.0444	0.4554	-0.2345	0.2260
Hightech indicator variable	0.1689	0.3747	0.0000	0.0000
Numerical Terms	83.3340	30.3325	60.8610	102.5450
Financial Statements	1.9013	0.9105	1.0000	3.0000
Total Words	888.1782	580.4106	483.0000	1151.0000
Log(Total Words)	6.5932	0.6448	6.1800	7.0484
PE Ratio	21.1502	47.6680	14.9225	35.14323
Forecast Dispersion	0.0569	1.0125	0.0256	0.1443
R&D Expenditures	0.0000	0.2960	0.0000	0.3721

Table 3. The Effect of Hard and Soft Information on Announcement Period CARs defined over [*t-1, t+1*]

In this table we present estimates of the following two equations:

$$\sum_{i=-1}^{1} AR_{jt+i} = \beta_0 + \beta_{SUE} SUE_{jt} + \beta_{Sent} SNetOpt_{jt} + \varepsilon_{jt},$$

$$\sum_{i=-1}^{+1} AR_{jt+i} = \beta_0 + \beta_{SUE} SUE_{jt} + \beta_{SUE\_Size} SUE_{jt} \times Size_{jt} + \beta_{SNO} SNetOpt_{jt}$$

$$+\beta_{SNO\_Size} SNetOpt_{jt} \times Size_{jt} + \varepsilon_{jt},$$

where SUE_{jt} is the standardized unexpected earnings, $SNetOpt_{jt}$, is the standardized unexpected net optimism in the earnings statement. The sample period includes all available earnings announcements from January, 1998 to July, 2006, for a total of 21,580 (13,907) firm-quarter observations for the Compustat (IBES) sample. We use standard errors clustered by calendar quarter and firm to compute the t-statistics and the p-values that are reported next to the coefficient estimates.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
	Panel A: Baseline Results					
Diction						
SUE	0.01051	13.19	0	0.02024	16.41	0
SNetOpt	0.00421	5.38	0	0.00201	2.77	0.0056
Intercept	0.00248	2.97	0.003	-0.00302	-2.65	0.008
Adjusted R-squared	1.72%			8.32%		
General Inquirer						
SUE	0.01020	12.72	0	0.02003	16.1	0
SNetOpt	0.00627	8.00	0	0.00349	3.79	0.0001
Intercept	0.00247	2.87	0.0042	-0.00299	-2.60	0.0093
Adjusted R-squared	2.00%			8.43%		
	Panel B: Baseline Results Controlling for Firm Size					
Diction						
SUE	0.06686	7.26	0	0.05407	3.93	0.0001
SUE×Size	-0.00285	-6.55	0	-0.00163	-2.65	0.0081
SNetOpt	0.02972	2.85	0.0043	0.02005	2.58	0.01
SNetOpt ×Size	-0.0013	-2.58	0.0098	-0.00088	-2.39	0.0167
Intercept	0.00281	3.37	0.0008	-0.00248	-2.33	0.0199
Adjusted R-squared	2.19%			8.52%		
General Inquirer						
SUE	0.06491	7.31	0	0.05174	3.71	0.0002
SUE×Size	-0.00277	-6.61	0	-0.00153	-2.45	0.0141
SNetOpt	0.03818	5.00	0	0.03281	3.6	0.0003
SNetOpt ×Size	-0.00163	-4.53	0	-0.00144	-3.43	0.0006
Intercept	0.00283	3.30	0.001	-0.00242	-2.26	0.0241
Adjusted R-squared	2.52%			8.68%		

Table 4. Announcement Period CARs with Firm Characteristics

In this table we present estimates of the following equation:

$$\sum_{i-1}^{+1} AR_{jt+i} = \beta_0 + \beta_{SUE} SUE_{jt} + \beta_{SUE\_Size} SUE_{jt} \times Size_{jt} + \beta_{SUE\_X} SUE_{jt} \times X_{jt}$$
$$+ \beta_{SNO} SNetOpt_{jt} + \beta_{SNO\_Size} SNetOpt_{jt} \times Size_{jt} + \beta_{SNO\_X} SNetOpt_{jt} \times X_{jt} + \varepsilon_{jt},$$

where SUE_{jt} is the standardized unexpected earnings, $SNetOpt_{jt}$, is the standardized unexpected net optimism in the earnings statement. The sample period includes all available earnings announcement dates from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. We use standard errors clustered by calendar quarter and firm to compute the t-statistics and the p-values that are reported next to the coefficient estimates.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
	Panel A: Analyst Coverage					
SUE	0.06873	6.65	0	0.04831	3.05	0.0023
SUE×Size	-0.00296	-5.87	0	-0.00131	-1.76	0.0783
SUE×Analyst Coverage	0.00007	0.66	0.5094	-0.00015	-1.12	0.2648
SNetOpt	0.03647	3.01	0.0026	0.0308	2.74	0.0061
SNetOpt×Size	-0.00169	-2.8	0.0052	-0.0015	-2.6	0.0093
SNetOpt× Analyst Cov.	0.00025	1.88	0.0603	0.00032	1.76	0.0788
Intercept	0.00283	3.37	0.0008	-0.00247	-2.31	0.0207
Adjusted R-squared	2.20%			8.53%		
	Panel B: Media Coverage					
SUE	0.08284	8.39	0	0.04799	3.66	0.0002
SUE×Size	-0.00367	-7.7	0	-0.0013	-2.13	0.0331
SUE×Media Coverage	0.00008	2.27	0.0233	-0.00004	-0.88	0.3764
SNetOpt	0.02844	2.14	0.0322	0.02004	1.96	0.0504
SNetOpt×Size	-0.00134	-2.03	0.0428	-0.00091	-1.90	0.0581
SNetOpt×Media Coverage	0.00008	1.72	0.0858	0.00002	0.44	0.6579
Intercept	0.00295	3.70	0.0002	-0.00227	-2.28	0.0229
Adjusted R-squared	2.64%			9.04%		
	Panel C: Past Turnover					
SUE	0.06679	7.17	0	0.05207	4.11	0
SUE×Size	-0.00284	-6.36	0	-0.00167	-2.72	0.0065
SUE×Turnover	-0.00007	-0.13	0.8986	0.00327	1.48	0.1397
SNetOpt	0.03008	2.91	0.0036	0.02038	2.61	0.009
SNetOpt×Size	-0.00136	-2.72	0.0064	-0.00099	-2.59	0.0097
SNetOpt×Turnover	0.00115	2.66	0.0078	0.00242	3.57	0.0004
Intercept	0.00281	3.35	0.0008	-0.00252	-2.31	0.0209
Adjusted R-squared	2.22%			8.88%		
	Panel D: Forecast Dispersion					
SUE				0.05533	3.94	0.0001
SUE×Size				-0.00168	-2.68	0.0074
SUE× Forecast Dispersion				-0.00105	-2.58	0.0099
SNetOpt				0.02181	2.77	0.0057
SNetOpt×Size				-0.00098	-2.64	0.0082
SNetOpt× Forecast Dispersion				0.00294	1.68	0.0929
Intercept				-0.00254	-2.43	0.0153
Adjusted R-squared				8.62%		

Table 4. Announcement Period CARs with Firm Characteristics (Continued)

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
	Panel E: Numerical Terms					
SUE	0.06939	7.41	0	0.05409	4.28	0
SUE×Size	-0.00272	-6.28	0	-0.00164	-2.51	0.0119
SUE×Numerical Terms	-0.00006	-3.24	0.0012	0.000001	0.02	0.9871
SNetOpt	0.0278	2.57	0.0101	0.01879	2.45	0.0143
SNetOpt×Size	-0.00132	-2.66	0.0078	-0.00087	-2.38	0.0171
SNetOpt×Numerical Terms	0.00003	1.82	0.0687	0.00001	0.69	0.4905
Intercept	0.00302	3.67	0.0002	-0.00245	-2.24	0.0248
Adjusted R-squared	2.24%			8.51%		
	Panel F: High Tech					
SUE	0.06661	7.63	0	0.0469	3.57	0.0004
SUE×Size	-0.00284	-6.78	0	-0.00138	-2.35	0.0189
SUE×High Tech	0.00043	0.15	0.8799	0.01013	3.48	0.0005
SNetOpt	0.02853	2.75	0.0059	0.01898	2.37	0.0176
SNetOpt×Size	-0.00128	-2.53	0.0114	-0.00088	-2.27	0.0231
SNetOpt×High Tech	0.00358	1.58	0.113	0.00371	1.73	0.0829
Intercept	0.00282	3.38	0.0007	-0.00263	-2.38	0.0173
Adjusted R-squared	2.21%			8.87%		
	Panel G: PE Ratio					
SUE				0.04564	3.32	0.0009
SUE×Size				-0.00126	-2.03	0.0428
SUE×PE Ratio				0.00001	0.96	0.3388
SNetOpt				0.01805	1.94	0.0526
SNetOpt×Size				-0.0005	-1.15	0.2487
SNetOpt×PE Ratio				0.00002	2.17	0.0304
Intercept				-0.00192	-1.61	0.1065
Adjusted R-squared				9.18%		
	Panel H: R&D Expenses					
SUE	0.06635	7.4	0	0.05364	3.91	0.0001
SUE×Size	-0.00284	-6.6	0	-0.00172	-2.81	0.005
SUE×R&D Expenses	0.00394	0.58	0.5638	0.02466	3.93	0.0001
SNetOpt	0.02942	2.86	0.0043	0.01965	2.47	0.0134
SNetOpt×Size	-0.00133	-2.67	0.0077	-0.00094	-2.47	0.0136
SNetOpt×R&D Expenses	0.01001	1.62	0.1058	0.01603	2.29	0.0222
Intercept	0.00282	3.39	0.0007	-0.00249	-2.34	0.0191
Adjusted R-squared	2.21%			8.81%		
	Panel I: EFKOS e-loading factor					
SUE	0.05971	6.85	0	0.04719	3.50	0.0005
SUE×Size	-0.0025	-6.10	0	-0.00131	-2.16	0.0305
SUE×EFKOS e-loading	0.00448	1.80	0.0725	0.00709	3.02	0.0025
SNetOpt	0.01885	1.94	0.0526	0.01223	1.75	0.0802
SNetOpt×Size	-0.00079	-1.52	0.1279	-0.00053	-1.32	0.1857
SNetOpt× EFKOS e-loading	0.00742	4.01	0.0001	0.00589	3.03	0.0024
Intercept	0.00288	3.63	0.0003	-0.00231	-2.20	0.0278
Adjusted R-squared	2.43%			8.88%		

Table 4. Announcement Period CARs with Firm Characteristics (Continued)

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
	Panel J					
SUE	0.06192	6.76	0	0.04598	3.72	0.0002
SUE×Size	-0.00237	-5.51	0	-0.00138	-2.07	0.0383
SUE×Numerical Terms	-0.00029	-2.92	0.0035	0.00002	0.08	0.9356
SUE×Turnover	-0.00013	-0.23	0.8143	0.00289	1.32	0.186
SUE× EFKOS e-loading	0.0043	1.73	0.0841	0.00584	2.55	0.0108
SNetOpt	0.01952	1.82	0.0681	0.01383	1.69	0.0902
SNetOpt ×Size	-0.00088	-1.73	0.0836	-0.00065	-1.66	0.0976
SNetOpt ×Numerical Terms	0.00017	1.89	0.0585	0.00009	0.92	0.3594
SNetOpt ×Turnover	0.00088	1.98	0.0478	0.00212	3.43	0.0006
SNetOpt × EFKOS e-loading	0.00709	3.71	0.0002	0.00509	2.67	0.0077
Intercept	0.00299	3.65	0.0003	-0.00252	-2.24	0.0253
Adjusted R-squared	2.45%			9.05%		

Table 5A. Long Horizon CARs defined over [t+2, t+62]

In this table we present estimates of the following equation:

$$\sum_{i=+2}^{62} AR_{jt+i} = \beta_0 + \beta_{SUE} SUE_{jt} + \beta_{Sent} SNetOpt_{jt} + \varepsilon_{jt},$$

where SUE_{jt} is the standardized unexpected earnings, $SNetOpt_{jt}$, is the standardized unexpected net optimism in the earnings statement. The sample period includes all available earnings announcements from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. We use standard errors clustered by calendar quarter and firm to compute the t-statistics and the p-values that are reported next to the coefficient estimates.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Diction						
SUE	0.0072	4.15	0	0.0041	2.68	0.0074
SNetOpt (Diction)	0.0027	1.85	0.065	0.0027	1.44	0.1493
Intercept	0.0004	0.21	0.8357	0.0018	0.89	0.3744
Adjusted R-squared	0.17%			0.10%		
General Inquirer						
SUE	0.0069	3.93	0.0001	0.0038	2.48	0.013
SNetOpt (GI)	0.0054	3.7	0.0002	0.0053	2.94	0.0032
Intercept	0.0004	0.2	0.8381	0.0018	0.92	0.3592
Adjusted R-squared	0.23%			0.17%		

Table 5B. Speed of Adjustment

In this table we present estimates of the following seemingly unrelated regression:

$$\sum_{i=+2}^{62} AR_{jt+i} = \beta_{10} + \beta_{1SUE} SUE_{jt} + \beta_{1Sent} SNetOpt_{jt} + \varepsilon_{1jt},$$

$$\sum_{i=-1}^{62} AR_{jt+i} = \beta_{20} + \beta_{2SUE} SUE_{jt} + \beta_{2Sent} SNetOpt_{jt} + \varepsilon_{2jt},$$

where SUE_{jt} is the standardized unexpected earnings, $SNetOpt_{jt}$ and is the standardized unexpected net optimism in the earnings statement. The sample period includes all available earnings announcements from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Diction						
SUE, β_{1SUE}	0.0072	5.68	0	0.0041	3.46	0.001
SNetOpt (Diction), β_{1Sent}	0.0027	2.07	0.039	0.0027	1.85	0.064
SUE, β_{2SUE}	0.0177	12.5	0	0.0243	18.15	0
SNetOpt (Diction), β_{2Sent}	0.0069	4.78	0	0.0047	2.88	0.004
$\frac{\beta_{1SUE}}{\beta_{2SUE}}$	0.4068			0.1687		
$\frac{\beta_{1Sent}}{\beta_{2Sent}}$	0.3913			0.5745		
$H_0: \frac{\beta_{1Sent}}{\beta_{2Sent}} = \frac{\beta_{1SUE}}{\beta_{2SUE}} (\chi^2 - stat)$		0.02	0.8745		6.10	0.0135
General Inquirer						
SUE, β_{1SUE}	0.0069	5.42	0	0.0038	3.17	0.002
SNetOpt (GI), β_{1Sent}	0.0054	4.18	0	0.0053	3.64	0
SUE, β_{2SUE}	0.0171	12.04	0	0.0238	17.69	0
SNetOpt (GI), β_{2Sent}	0.0116	8.15	0	0.0088	5.38	0
$\frac{\beta_{1SUE}}{\beta_{2SUE}}$	0.4035			0.1597		
$\frac{\beta_{1Sent}}{\beta_{2Sent}}$	0.4655			0.6023		
$H_0: \frac{\beta_{1Sent}}{\beta_{2Sent}} = \frac{\beta_{1SUE}}{\beta_{2SUE}} (\chi^2 - stat)$		0.48	0.487		22.24	0

Table 5C. Long Horizon CARs defined over [t+2, t+62] and Certainty

In this table we present estimates of the following equation:

$$\sum_{i=+2}^{62} AR_{jt+i} = \beta_0 + \beta_{SUE} SUE_{jt} + \beta_{SUE\_Size} SUE_{jt} \times Size_{jt} + \beta_{SUE\_Cert} SUE_{jt} \times \text{Certainty}_{jt}$$
$$+ \beta_{SNO} SNetOpt_{jt} + \beta_{SNO\_Size} SNetOpt_{jt} \times Size_{jt} + \beta_{SNO\_Cert} SNetOpt_{jt} \times \text{Certainty}_{jt} + \varepsilon_{jt},$$

where SUE_{jt} is the standardized unexpected earnings, $SNetOpt_{jt}$, is the standardized unexpected net optimism in the earnings statement. The sample period includes all available earnings announcements from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. We use standard errors clustered by calendar quarter and firm to compute the t-statistics and the p-values that are reported next to the coefficient estimates.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Diction						
SUE	0.01825	0.99	0.3242	0.02513	1.14	0.2552
SUE×Size	-0.00044	-0.52	0.6042	-0.00117	-1.19	0.2347
SUE×Certainty	-0.00088	-0.59	0.5572	0.00133	1.16	0.2447
SNetOpt	0.01918	1.14	0.2529	0.00234	0.12	0.9066
SNetOpt×Size	-0.0005	-0.60	0.5457	0.00054	0.59	0.5566
SNetOpt× Certainty	-0.00272	-2.44	0.0147	-0.0043	-2.31	0.0207
Intercept	0.00039	0.2	0.8389	0.00198	1.03	0.3033
Adjusted R-squared	0.18%			0.16%		
General Inquirer						
SUE	0.01514	0.81	0.417	0.0206	0.93	0.3543
SUE×Size	-0.00031	-0.37	0.7139	-0.00097	-0.98	0.3283
SUE×Certainty	-0.0008	-0.53	0.5958	0.00131	1.10	0.2705
SNetOpt	0.04509	2.92	0.0035	0.04547	2.17	0.0299
SNetOpt×Size	-0.00166	-2.25	0.0247	-0.00156	-1.65	0.0984
SNetOpt× Certainty	-0.00271	-2.64	0.0083	-0.00321	-1.76	0.0777
Intercept	0.00041	0.21	0.8308	0.00211	1.10	0.2723
Adjusted R-squared	0.28%			0.24%		

Table 6A. Post-Announcement Drift

Each calendar quarter stocks are classified in one of three groups according to their earnings announcement surprise terciles. The surprise tercile for firm *i* in quarter *t* is a ranking from 1 to 3 of the earnings surprise, SUE_{it}, based on the previous quarter's surprise tercile cutoffs. The 3-day abnormal return is the cumulative size and B/M adjusted return over trading days [-1,+1], where day 0 is the earnings announcement date. The 60-day abnormal return after the announcement is the cumulative size and B/M adjusted return over trading days [+2, +62]. The cumulative returns are multiplied by 100. Firms in market capitalization deciles 9 and 10 are assigned to the large-firm group, firms in deciles 6 through 8 are assigned to the medium-firm group and those in deciles 1 to 5 are assigned to the small-firm group. Three, two and one asterisk denote, respectively, that the estimates are statistically significant at the one, five and ten percent level.

Tercile	SUE	3-Day CAR	60-Day CAR	3-1
		Small		
1	-0.961***	-1.847***	-1.998***	
2	0.065***	-0.112	-0.274	
3	1.075***	2.857***	1.017***	3.015***
		Medium		
1	-0.986***	-0.475***	0.161	
2	0.078***	0.269*	0.903***	
3	1.066***	1.45***	1.845***	1.683***
		Large		
1	-0.938***	-0.159	-0.393	
2	0.078***	0.346**	-0.056	
3	0.974***	1.155***	0.327	0.719

Table 6B. Post-Announcement Drift Revisited

Each calendar quarter stocks are classified in one of three groups according to their earnings announcement surprise terciles and net optimism surprise terciles. The surprise tercile for firm *i* in quarter *t* is a ranking from 1 to 3 of the earnings surprise, SUE_{it} , and net optimism surprise, $Sent_{it}$, based on the previous quarter's surprise tercile cutoffs. We label tercile 1 if both surprises fall in the first tercile, tercile 2 if both surprises fall in the second tercile and tercile 3 if both surprises fall in the third tercile. The 3-day abnormal return is the cumulative size and B/M adjusted return over trading days [-1,+1], where day 0 is the earnings announcement date. The 60-day abnormal return after the announcement is the cumulative size and B/M adjusted return over trading days [+2, +62]. The cumulative returns are multiplied by 100. Firms in market capitalization deciles 9 and 10 are assigned to the large-firm group, firms in deciles 6 through 8 are assigned to the medium-firm group and those in deciles 1 to 5 are assigned to the small-firm group. Three, two and one asterisks denote, respectively, that the estimates are statistically significant at the one, five and ten percent level.

Tercile	SUE	ΔNetOpt	3-Day CAR	60-Day CAR	3-1
			Small		
1	-0.999***	-1.104***	-2.436***	-2.217***	
2	0.068***	0.005	-0.079	0.185	
3	1.085***	1.035***	3.956***	1.885***	4.102***
			Medium		
1	-1.049***	-1.041***	-1.053***	0.745	
2	0.077***	0.016**	0.238	0.938**	
3	1.079***	1.014***	1.701***	2.84***	2.095**
			Large		
1	-0.981***	-0.995***	-0.21	-0.872	
2	0.081***	0.004	0.313	-0.074	
3	0.94***	0.986***	1.259***	-0.042	0.830

Table 7A: Announcement period volatility defined over [t-1, t+1]

In this table we present estimates of the following equation:

$$\log(\sum_{i=-1}^{1} AR_{jt+i}^{2}) = \gamma_0 + \gamma_1 \log(\sum_{i=2}^{4} AR_{jt-i}^{2}) + \gamma_2 \left|SUE_{jt}\right| + \gamma_3 \left|SUE_{jt}\right| \times I(SUE_{jt} < 0)$$

$$+\gamma_4 \left|SNetOpt_{jt}\right| + \gamma_5 \left|SNetOpt_{jt}\right| \times I(SNetOpt_{jt} < 0) + \gamma_6 certainty_{jt} + \overline{\gamma}_5 \overline{\Upsilon}_{jt}$$

We measure the volatility of abnormal returns during the event window as the logarithm of the sum of the absolute value of abnormal returns during the [t-1, t+1] event window. The rest of the variables are defined in Appendix B. The sample period includes all available earnings announcement dates from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. We use standard errors clustered by calendar quarter and firm to compute the t-statistics and the p-values that are reported next to the coefficient estimates.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Past Volatility	0.39897	34.51	0	0.37541	29.81	0
|SUE|	0.08127	4.98	0	0.20903	13.36	0
|SUE|×I(SUE<0)	-0.0452	-2.00	0.046	-0.00301	-0.12	0.9071
|SNetOpt|	0.0777	3.21	0.0013	0.02593	0.88	0.3803
|SNetOpt|×I(SNetOpt<0)	0.00982	0.38	0.705	0.01903	0.65	0.515
Certainty	-0.14281	-11.06	0	-0.133	-9.21	0
Adjusted R-squared	15.38%			14.10%		
Past Volatility	0.28581	24.77	0	0.24792	17.66	0
|SUE|	0.11918	6.99	0	0.23951	14.51	0
|SUE|×I(SUE<0)	-0.05025	-2.11	0.0349	-0.00837	-0.34	0.7329
|SNetOpt|	0.04979	2.5	0.0124	0.00034	0.02	0.9867
|SNetOpt|×I(SNetOpt<0)	0.0035	0.14	0.8853	0.01704	0.59	0.5545
Certainty	-0.05734	-4.88	0	-0.042	-3.27	0.0011
Log(Market Cap.)	-0.07666	-6.62	0	-0.10295	-6.47	0
Log(1+analyst)	0.13037	7.88	0	0.19962	6.78	0
EFKOS e-Loading	0.24387	6.61	0	0.27088	6.81	0
REG_FD	-0.14345	-1.53	0.126	-0.17415	-1.47	0.141
Time Trend	-0.00241	-0.66	0.5083	-0.00403	-0.87	0.3859
Financial Statements	0.06254	3.57	0.0004	0.07612	3.71	0.0002
Log(Total Words)	-0.04818	-2.36	0.0184	-0.04788	-1.77	0.0762
Recent Media Coverage	0.03182	3.84	0.0001	0.01807	2.03	0.0422
Log(MB Ratio)	0.108	5.49	0	0.15965	6.53	0
-1/(1+age)	0.1358	0.61	0.5406	-0.07147	-0.3	0.7621
I(Dividend Payout)	-0.68934	-19.44	0	-0.68412	-17.59	0
Leverage	-0.48688	-5.74	0	-0.56016	-5.55	0
Forecast Dispersion				0.02362	0.5	0.6189
Adjusted R-squared	22.63%			22.93%		

Table 7B: Announcement period volatility defined over [t-1, t+1], Extended Model

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Past Volatility	0.28314	24.44	0	0.24517	17.46	0
|SUE|	0.11341	6.51	0	0.22801	13.55	0
|SUE|×I(SUE<0)	-0.0605	-2.58	0.0098	-0.00398	-0.16	0.8713
|SNetOpt|	0.04181	2.15	0.0313	-0.00551	-0.28	0.7786
|SNetOpt|×I(SNetOpt<0)	0.00002	0	0.9992	0.01436	0.5	0.6162
Certainty	-0.03548	-3.13	0.0018	-0.02582	-1.98	0.0474
Log(Market Cap.)	-0.10574	-9.29	0	-0.13376	-7.83	0
Log(1+analyst)	0.1348	8.2	0	0.20716	6.96	0
EFKOS e-Loading	0.24158	6.72	0	0.2715	7.01	0
REG_FD	-0.11368	-1.23	0.218	-0.15037	-1.24	0.2136
Time Trend	-0.00323	-0.91	0.3633	-0.00467	-1.01	0.3143
Financial Statements	0.06252	3.7	0.0002	0.07778	3.93	0.0001
Log(Total Words)	-0.01649	-0.86	0.3898	-0.01575	-0.6	0.5465
Recent Media Coverage	0.02864	3.54	0.0004	0.01459	1.69	0.0907
Log(MB Ratio)	0.15228	7.05	0	0.20319	7.6	0
-1/(1+age)	0.04502	0.19	0.8467	0.05245	0.22	0.829
I(Dividend Payout)	-0.6145	-17.71	0	-0.62092	-16.35	0
Leverage	-0.53515	-6.83	0	-0.60016	-6.19	0
ROE	-1.14771	-2.93	0.0034	-0.91165	-2.2	0.0281
Vol(ROE)	4.04076	12.24	0	3.66049	9.82	0
Forecast Dispersion				0.00538	0.11	0.9103
Adjusted R-squared	23.61%			23.75%		

Table 8A: Long horizon volatility defined over [t+2, t+62]

In this table we present estimates of the following equation:

$$\log(\sum_{i=+2}^{62} AR_{jt+i}^2) = \gamma_0 + \gamma_1 \log(\sum_{i=2}^{62} AR_{jt-i}^2) + \gamma_2 \left|SUE_{jt}\right| + \gamma_3 \left|SUE_{jt}\right| \times I(SUE_{jt} < 0)$$

$$+\gamma_4 \left|SNetOpt_{jt}\right| + \gamma_5 \left|SNetOpt_{jt}\right| \times I(SNetOpt_{jt} < 0) + \gamma_6 certa\operatorname{int} y_{jt} + \overline{\gamma}_5 \overline{\mathrm{Y}}_{jt}$$

We measure the volatility of abnormal returns during the event window as the logarithm of the sum of the absolute value of abnormal returns during the [t-1, t+1] event window. The rest of the variables are defined in Appendix B. The sample period includes all available earnings announcement dates from January, 1998 to July, 2006, and in total there are 21,580 (13,907) firm-quarter observations in the Compustat (IBES) sample. We use standard errors clustered by calendar quarter and firm to compute the t-statistics and the p-values that are reported next to the coefficient estimates.

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Past Volatility	0.79748	63.24	0	0.79545	58.26	0
|SUE|	0.0034	0.28	0.7777	-0.02315	-1.68	0.0925
|SUE|×I(SUE<0)	0.00768	0.73	0.4659	0.04253	2.56	0.0105
|SNetOpt|	0.02899	3.17	0.0015	0.03005	2.52	0.0119
|SNetOpt|×I(SNetOpt<0)	-0.01613	-1.38	0.1671	-0.02623	-1.87	0.0618
Certainty	-0.02106	-2.24	0.0248	-0.0289	-3.04	0.0024
Adjusted R-squared	63.34%			62.93%		
Past Volatility	0.57211	44.54	0	0.57002	37.73	0
|SUE|	0.02977	2.66	0.0079	0.01969	1.87	0.0618
|SUE|×I(SUE<0)	-0.00813	-0.89	0.3709	0.00961	0.76	0.4479
|SNetOpt|	0.02552	3.23	0.0012	0.02405	2.53	0.0116
|SNetOpt|×I(SNetOpt<0)	-0.01638	-1.66	0.0963	-0.02735	-2.22	0.0266
Certainty	-0.01991	-4.22	0	-0.02182	-3.45	0.0006
Log(Market Cap.)	-0.1094	-13.57	0	-0.10742	-10.34	0
Log(1+analyst)	0.01812	2.36	0.0184	0.03932	3.2	0.0014
EFKOS e-Loading	0.09416	5.28	0	0.08266	4.72	0
REG_FD	-0.09493	-0.93	0.3513	-0.1277	-1.19	0.2351
Time Trend	-0.01629	-3.63	0.0003	-0.01718	-3.53	0.0004
Financial Statements	0.00277	0.55	0.5794	0.00807	1.16	0.2476
Log(Total Words)	0.01066	0.99	0.3234	0.02131	1.67	0.0949
Recent Media Coverage	0.02224	6.08	0	0.02045	6.04	0
Log(MB Ratio)	0.06952	5.91	0	0.07588	6.09	0
-1/(1+age)	-0.07081	-0.65	0.5189	-0.18059	-1.73	0.0834
I(Dividend Payout)	-0.26855	-14.72	0	-0.25344	-12.29	0
Leverage	-0.12904	-4.27	0	-0.09954	-2.56	0.0106
Forecast Dispersion				0.02811	1.26	0.2086
Adjusted R-squared	68%			67.27%		

Table 8B: Long horizon volatility defined over [t+2, t+62], Extended Model

	Compustat Sample			IBES Sample		
	coefficient	t-stat	p-value	coefficient	t-stat	p-value
Past Volatility	0.55652	42.18	0	0.5526	36.05	0
|SUE|	0.03386	3.06	0.0022	0.02079	2.06	0.0398
|SUE|×I(SUE<0)	-0.02447	-2.71	0.0066	0.00072	0.06	0.9519
|SNetOpt|	0.02325	3	0.0027	0.02281	2.45	0.0142
|SNetOpt|×I(SNetOpt<0)	-0.01649	-1.71	0.0881	-0.02762	-2.27	0.0231
Certainty	-0.01181	-2.39	0.0168	-0.01396	-2.17	0.0298
Log(Market Cap.)	-0.11142	-14.71	0	-0.11023	-11.07	0
Log(1+analyst)	0.02046	2.61	0.009	0.04308	3.42	0.0006
EFKOS e-Loading	0.07909	4.58	0	0.06826	3.96	0.0001
REG_FD	-0.09138	-0.91	0.3618	-0.12416	-1.16	0.2476
Time Trend	-0.01689	-3.78	0.0002	-0.01803	-3.66	0.0002
Financial Statements	0.00268	0.53	0.5963	0.01075	1.52	0.1293
Log(Total Words)	0.0151	1.43	0.1534	0.02915	2.21	0.0273
Recent Media Coverage	0.01922	5.3	0	0.01657	4.96	0
Log(MB Ratio)	0.06354	5.67	0	0.07383	6.63	0
-1/(1+age)	-0.08045	-0.73	0.4662	-0.13959	-1.27	0.2035
I(Dividend Payout)	-0.24548	-14.98	0	-0.22851	-12.15	0
Leverage	-0.17691	-5.87	0	-0.15678	-4	0.0001
ROE	-1.77087	-10.82	0	-2.05148	-8.84	0
Vol(ROE)	0.6427	2.85	0.0043	0.72795	2.79	0.0053
Forecast Dispersion				0.01446	0.65	0.5185
Adjusted R-squared	68.34%			67.69%		