Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

News versus Sentiment: Predicting Stock Returns from News Stories

Steven L. Heston and Nitish R. Sinha

2016-048

Please cite this paper as:
Heston, Steven L. and Nitish R. Sinha (2016). "News versus Sentiment: Predicting Stock Returns from News Stories," Finance and Economics Discussion Series 2016-048. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2016.048.

News versus Sentiment: Predicting Stock Returns from News Stories

June 6, 2016

Abstract

This paper uses a dataset of more than 900,000 news stories to test whether news can predict stock returns. We measure sentiment with a proprietary Thomson-Reuters neural network. We find that daily news predicts stock returns for only 1 to 2 days, confirming previous research. Weekly news, however, predicts stock returns for one quarter. Positive news stories increase stock returns quickly, but negative stories have a long-delayed reaction. Much of the delayed response to news occurs around the subsequent earnings announcement.

Steven L. Heston and Nitish Ranjan Sinha. Heston: Department of Finance, Robert H. Smith School of Business, University of Maryland, College Park, *sheston@rhsmith.umd.edu*. Ranjan Sinha: Board of Governors of the Federal Reserve System, *nitish.r.sinha@frb.gov*. We thank Tim Loughran, Paul Tetlock, seminar participants at University of Maryland (finance brownbag), Acadian Asset Management and the Office of Financial Research. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors.

JEL–Classification: G12, G14

Keywords: News, Text Analysis

1 Introduction

Textual information processing has become a growing part of financial practice. Duhigg (2006) and Ro (2012) write about general artificial intelligence for stock picking, while Lo (1994) reviews neural networks. Specific applications include bankruptcy prediction Atiya (2001), corporate distress diagnosis Altman, Marco, and Varetto (1994), and consumer credit risk Khandani, Kim, and Lo (2010). While industry has led the applications, academic empirical research is increasingly confirming the value of textual analysis. Tetlock's pioneering studies ((Tetlock, Saar-Tsechansky, and Macskassy 2008) and (Tetlock 2007)) demonstrate that news stories contain information relevant to predicting both earnings and stock returns. Subsequent studies have applied similar techniques with a variety of news sources. Researchers have generally found that textual information can briefly predict returns at the aggregate market level ((Tetlock 2007), (Dougal, Engelberg, García, and Parsons 2012), (Garcia 2013) and Dzielinski and Hasseltoft (2013)) as well at the individual stock level ((Boudoukh, Feldman, Kogan, and Richardson 2013), (Sinha 2016) and (Chen, De, Hu, and Hwang 2014)). However, the research has been limited to a comparatively narrow event window, and has not shown significant predictability beyond two days after news release. In contrast, this paper uses a neural network to show that news stories can predict stock returns for up to 13 weeks.

The rapid growth of this empirical research has entailed the use of different datasets and methodologies. For example, Tetlock, Saar-Tsechansky, and Macskassy (2008) uses a broad sample of *Wall Street Journal* and Dow Jones News Service articles, whereas Loughran and McDonald (2011) use more specialized 10-K filings. Similarly, Garcia (2013) analyzes New York Times articles, whereas

Jegadeesh and Wu (2013) also examine 10-K's, Lerman and Livnat (2010) uses 8-K's, and Chen et al (2014) use social media. These conflicting choices confound the type of source documents used for the textual analysis with the type of textual processing. In particular, it begs the question of whether textual processing can effectively predict stock returns based on a broad set of text sources.

In addition to methodological differences, empirical studies have found different types of predictability in applications at the aggregate market level or the individual stock level. Early work by Tetlock (2007) finds that short-term return predictability is quickly reversed at the market level. Loughran and McDonald (2011) find greater response for individual stocks within a multi-day event window. Garcia (2013) and Jegadeesh and Wu (2013) also find different results with market returns and individual stocks, respectively. More recently, Hillert, Jacobs, and Müller (2014) suggest that media overreaction underlies stock momentum. Hagenau, Hauser, Liebmann, and Neumann (2013) measure news momentum to predict CDAX index returns, and Uhl, Pedersen, and Malitius (2015) aggregate sentiment for tactical asset allocation. In addition to aggregate market returns versus individual stocks, differences might stem from different source of text, or different methodologies. The duration and reversal of return predictability are important because the economic interpretation of news depends on whether there is a permanent news impact or a transient impact. Permanent news impact would suggests news as information on the other hand transient news impact would suggest news as sentiment. As Tetlock (2007) summarizes, "The sentiment theory predicts short-horizon returns will be reversed in the long run, whereas the information theory predicts they will persist indefinitely."

This paper examines stock return predictability using a sophisticated neural network.[1] It applies these techniques on a large common set of Reuters news releases. We find that the neural network appears to extract permanent information that is not fully impounded into current stock prices.

The duration of return predictability depends critically on the portfolio formation procedure. Previous research by Tetlock, Saar-Tsechansky, and Macskassy (2008), Loughran and McDonald (2011), and Lerman and Livnat (2010) has established a short-term response of stock prices to news. We also find that stocks with positive (negative) news over one day have subsequent predictably high (low) returns for 1 to 2 days. But going beyond the published literature, we find that aggregating news over one week produces a dramatic increase in predictability of returns. Stocks with news over the past week have predictable returns for up to 13 weeks, which is true even for stocks with only one news event per week. The difference in return predictability depending on the aggregation horizon shows that it is important to gauge relative news sentiment by examining news over longer horizon rather than just one day of stories.

Our study controls for neutral news stories to isolate the effect of news tone on stock returns. Controlling for neutral news is essential to distinguish a publication effect from an informative news effect. We confirm the finding of Fang and Peress (2009) that firms without news have different returns than firms with news. If no-news firms are compared to firms with news, then this effect can distort the comparison of positive news with negative news. Instead, we control for the effect of positive and negative news by comparing with neutral news. We find that news tone does indeed have an effect on stock returns. Positive news only predicts positive returns for about one week, but negative news predicts negative

[1] Antweiler and Frank (2004) use Naive Bayes classifier to classify text. Das and Chen (2007) examines the effect of chat board messages on the stock prices using a voting scheme across multiple classifiers.

returns for up to a quarter. Reaction to negative news over a longer horizon suggests that short sale constraints might slow the incorporation of information extracted by our textual processing techniques.

Section 2 describes the data and textual processing methods we use. Section 3 compares the ability of these textual processing methods to predict stock returns. Section 4 shows the existence of a distinct news effect, and controls for this effect to contrast the different predictive pattern of positive and negative news sentiment. It also describes the pattern of return predictability around future earnings announcements. A final section concludes.

2 Textual Processing

The primary purpose of our study is to forecast individual stock returns using textual analysis of news stories based on a neural network. Internet news sources and social media are providing a growing universe of textual information, including internet searches (Da, Engelberg, and Gao 2015), Facebook networks (Simon and Heimer 2012), and Twitter broadcasts (Bollen, Mao, and Zeng 2011). Analysis of these sources typically requires complex analytic tools, which give potential power to predict returns but also makes the analysis inherently opaque.[2] Therefore, we perform diagnostics to find patterns that suggest economic reasons for predictability. A distinguishing feature of our analysis is a broad dataset of news items. For example, Tetlock (2007) analyzes the *Wall Street Journal's* "Abreast of the Market" column, and Tetlock, Saar-Tsechansky, and Macskassy (2008) extended the analysis to firm-specific stories in the *Wall Street Journal* and the Dow-Jones News Service. Loughran and McDonald (2011) use a more specialized

[2]Butler (2013) criticize this lack of transparency and associated interpretation problems when diagnosing the ability of Google searches to forecast influenza outbreaks.

list of financial words to analyze company 10-K reports. Our analysis addresses the question of whether the improvement in results from specialized processing persists in a broad dataset or whether it requires suitably specialized textual input.

Another motivation for using a large, broad dataset is to increase the power to distinguish different types of return predictability. Temporary market sentiment or news-induced trading liquidity should be quickly reversed. Boudoukh, Feldman, Kogan, and Richardson (2013) predict that markets will overreact to simple news, and underreact to complex news. In particular, complex new information should have a permanent impact on stock prices. Larger datasets and more powerful textual analysis methods have the potential to detect these distinct patterns of predictability. For example, we find that weekly news predicts returns much longer than daily news.

Our dataset has a measure of the "tone" or sentiment of each news story.[3] The story-specific sentiment measure allows us to distinguish the effect of news publication from the effect of favorable or unfavorable news. The publication of news may draw attention to a stock, inducing both rational and irrational trading. This may affect the liquidity of the stock, and consequently change the expected return. We show that stocks with news have different expected returns from stocks without news. Controlling for this publication effect shows that positive and negative news are incorporated into stock prices at different speeds.

Our empirical analysis uses 900,754 articles tagged with firm identifiers from the Thomson-Reuters news system over the calendar years 2003 to 2010. Thomson Reuters provides a dataset of news sentiment called Thomson Reuters NewS-

[3]By contrast, Akbas, Boehmer, Erturk, and Sorescu (2013) use the observed stock return to classify the tone of an article. The Ravenpack database used by von Beschwitz, Keim, and Massa (2013) also uses sentiment analytics.

cope Data (sentiment data). The dataset is broader and larger than many of the datasets previously studied.[4] The dataset identifies the time of the news story (with millisecond resolution), the firm mentioned in the story, the headline of the news story, the story ID, the relevance of the news article for the firm, the staleness of a news item, and measures from a neural-network-based sentiment engine. Thomson Reuters also provides a dataset called the Thomson Reuters news archive (text data), which contains the time of the news story, the story ID, the headline of the news story, and the full text of the news item. We match the sentiment data with the text data using the timestamp and story ID for all the items in the sentiment data and obtain a dataset that contains the text as well as the respective probabilities for the article being positive, negative, and neutral. We exclude news items linked to more than one article in the sample, to ensure that this information did not appear in the sample before. We also exclude news about firms that could not match to any ticker symbol in the CRSP dataset and articles about firms with relevance scores below 35%.[5]

These stories are tagged by Thomson Reuters with several topic codes. The appendix lists all topic codes with a brief descriptions and the proportion of news articles being tagged with a particular topic code. The three most commonly used topic codes are 'STX', 'RES', and 'MRG'. The topic code 'STX' indicates additions and deletions from stock indices, new listings, delistings and suspensions; it has been assigned to 13% of news articles in our sample. The topic code 'RES' indicates all corporate financial results, tabular and textual reports, dividends, annual and quarterly reports; it has been assigned to 14% of news articles in

[4]The Thomson-Reuters sentiment dataset has also been used by Riordan, Storkenmaier, Wagener, and Sarah Zhang (2013), (Leinweber and Sisk 2011) and (Healy and Lo 2011), among others.

[5]Boudoukh et al (2013) show that relevant news affects stock returns more than irrelevant news.

our sample. The topic code 'MRG' indicates mergers and acquisitions; it has been assigned to 18% of news articles. Most of remaining topic codes indicate economic news.

The Thomson Reuters sentiment engine first identifies parts of speech and morphologically stems the words by matching each word to its root word. For example, "gone," "went," and "goes" are all identified as "go." The sentiment engine does shallow parsing whereby it identifies the subject of the sentence, and then identifies words as adjectives, adverbs, intensifiers, nouns, and verbs. This lexical identification is important for sentiment processing because certain phrases and parts of speech tend to convey tone. The lexical identification also recognizes negation, intensification, and verb resolution. The final sentiment classification uses these features as inputs to a three-layer back-propagation neural network classifier. The classifier was trained using a random sample of 3,000 triple-annotated news articles spanning 14 months from December 2004 to January 2006. Analysts who analyze blogs and other outlets of public opinion annotated the news articles. The annotation order was randomized so that manual annotators would not have been able to anticipate stock returns from reading the news articles. Given that the training sample was less than 1% of our data, the effect of data-snooping is miniscule.[6] The engine is described in greater detail in Sinha (2016) and Infonic (2008).

===Insert Table 1 here====

Table 1 shows the average net sentiment (positive minus negative) is slightly positive, at 2.4%. But Table 2 shows news is not uniformly distributed across firms.

[6]Informal inspection of results indicates no significant difference in predictability in the post-training period.

===Insert Table 2 here====

Table 2 shows summary statistics for firms sorted by size. Firms in the largest decile have frequent stories; 22.42 stories per week which exceeds three stories per day. But the small firms are comparatively neglected in news coverage. Firms in the smallest two deciles average less than one article per week, and more than 90% of firms in the smallest three size deciles receive no news in a given week. The greater news coverage of large firms ensures that small firms will not dominate our news-based return strategies. This alleviates concerns that profits are associated with exposure to illiquidity. Over the 10-year sample period, small firms underperformed large firms; the smallest decile lost 0.14% per week, while the largest decile gained 0.06% per week. The last column reveals a critical feature relevant to our study of textual analysis. Firms that receive news coverage in a given week have different average returns than typical firms of their size in the subsequent week. Because these "no news" returns occur in the ensuing weeks, they are not subject to a survivorship bias or short-term informational effect, as in "the dog that didn't bark." The most dramatic difference occurs in the smallest decile, where the average small firm lost 0.14% in a given week, but small firms with news averaged 2.00% in the week following the news. Given that small firms are often illiquid and costly to trade, the return differential between small firms with and without news may not represent a profit opportunity. But it documents that firms with news are distinctly different from firms without news. When measuring the effect of news sentiment, it is important to control for the existence of news.

3 Predicting Returns

Our simplest test of news sentiment uses portfolios based on net sentiment, positive minus negative. In contrast to previous studies that use SEC filings or periodic newspaper columns, our dataset has almost one million news stories, sometimes with multiple stories about a particular firm. Therefore, we measure the sentiment for a given firm as the average positive minus negative sentiment on all stories about that firm in a formation period. Table 3 presents excess returns on quintile spreads, i.e., the difference between returns on the highest and lowest sentiment portfolios. The quintiles are formed daily on Day 0, and returns are reported daily. We use quintiles instead of deciles or more selective portfolios because many stocks do not have news results on any given day. Note that we can interpret these excess returns as differences between returns in excess of a benchmark market portfolio, consistent with the methodology of Brown and Warner (1985). In particular, these excess returns difference out components due to the risk-free rate or market return.

===Insert Table 3 here====

The contemporaneous returns on the Day 0 news release show economically and statistically significant announcement day returns of 1.99%. This is quite large for a single day return, and shows the impact of news on stock prices. Note that average excess returns on the quintile spreads are invariably positive in the 10 days preceding the publication of news, usually with t-statistics exceeding 2. This is expected, since news stories may lag events that affect stock prices. It suggests that stock returns predict news, rather than the converse.

The more interesting result is the post-publication returns. The neural network produces returns of 0.17% on Day 1 and 0.04% on Day 2, both significant at the 95% level. It appears that this method of textual processing predict stock returns that are not immediately reversed. In particular, these returns do not appear to be an artifact of bid-ask bounce or temporary liquidity imbalances.

===Insert Figure 1 here====

Figure 1 contrasts these methods by showing cumulative daily excess returns for one quarter, or 63 business days, after the news date. The subsequent performance is rather flat, suggesting that information is quickly absorbed into prices.

These daily results are consistent with previous findings with daily data. Tetlock, Saar-Tsechansky, and Macskassy (2008), Loughran and McDonald (2011), and Lerman and Livnat (2010) all find predictability over event windows of 1 to 4 days, with varying degrees of reversal. This previous research used periodic news columns and SEC filings. Those datasets typically have only one news item per firm. In contrast, our dataset often has multiple news stories about firms spread over several adjacent days. Given our dataset with frequent stories, daily aggregation might not be the best choice.

===Insert Table 4 here====

Table 4, Panel A shows that the predictability changes dramatically when decile portfolios are formed based on weekly news.[7] The announcement week decile spread produces an excess return of 3.75%. This number must be interpreted with caution, since the news stories may be published subsequent to a day on which news actually caused high returns. But the subsequent weekly returns are truly out-of-sample. The neural network predicts subsequent returns for 13

[7]We found similar results with biweekly and monthly aggregation.

weeks after the new story release. Most of them are statistically significant at the 95% level, including a 0.21% return in Week 13. Figure 2 graphs the cumulative returns from the weekly strategy. In contrast to the daily results of Figure 1, it shows a persistent upward trend of profitability.

===Insert Figure 2 here====

It is conceivable that the large posts-news returns are compensation for exposure to firm risk or characteristics. Two likely candidates are size and momentum. As Table 2 shows, size is clearly negatively related to the volume of news stories. To the extent that size represents exposure to risk, or proxies for some other anomalous return factors, it is useful to control for it. Momentum is an even more related factor, because we have already shown that good news tends to be preceded by positive returns. Hence, the returns due to news might be a by-product of Jegadeesh and Titman (1993) momentum factor. Therefore, Table 4 presents two additional columns that control for size and momentum, respectively. In each week, we assign firms to deciles based on their size or momentum. Then instead of using firm returns, we use returns in excess of their size or momentum categories. The results show that size and momentum not subsume the return predictability of news.

This still leaves open the question of why weekly news formation predicts returns for 13 weeks, while daily formation predicts returns for only two days. There are two explanations for the striking improvement in predictability when using weekly returns. One explanation is that some firms have multiple news stories over different days within a week, and the predictability stems from the information confirmation of these clustered news stories. Of the firms with news in a given week, only 35% have more than one news story, and only 9% have more than two. These firms with multiple news stories tend to be larger than firms with less news coverage. This explanation argues that this minority of firms drives the profitability of weekly strategies.

===Insert Figure 3 here====

A second, more prosaic explanation is that the distribution of daily news is quite variable over time. Figure 3 illustrates the higher volatility for daily news sentiment by graphing the 20th and 80th percentiles of Thomson-Reuters news sentiment based on daily and weekly news. It is clear that the thresholds for daily quintile sentiment are quite volatile.[8] Some days simply have little news, or little news with strong sentiment, while others have an abundance of news with strong positive or negative sentiment. At a few points, the daily 20th and 80th percentile lines almost touch. The small difference between the 20th and 80th percentile on such days means that firms with stories in the highest quintile of sentiment on one day would be in the lowest quintile on an adjacent day. Clearly, daily news sentiment is a noisy way of classifying firms based on sentiment. The weekly cutoffs still show some variation but are much more stable over time.

===Insert Figure 4 here====

Table 4, Panel B reports the weekly decile returns for subsamples of firms that have one and multiple news stories in Week 0. Both subsamples are profitable in 12 of the 13 post-news weeks. The decile spread of firms with multiple news stories is positive at the 95% level of significance in Week 11, whereas the decile spread of firms with a single weekly news story is significantly positive at the 95% level in Week 13. Figure 4 graphs the cumulative returns to these strategies over

[8]One can also see the news thresholds shrank after the Thomson-Reuters database was expanded in 2005.

13 weeks. It shows that the decile strategy based on firms with multiple weekly news items is more profitable over the quarter, but both subsamples generate a strong, profitable trend. These positive decile spreads based on weekly news are quite different from the flat daily results in Figure 1.

4 News, Sentiment, and Earnings

The previous results show a delayed reaction to weekly news about firms. Specifically, firms with good news over a one-week period subsequently outperform firms with bad news over a one-week period. Portfolios formed on this basis earn excess returns for up to 13 weeks.

However, this exercise does not completely disentangle the "news effect" from the "sentiment effect". It is conceivable that the mere publication of news about a firm affects its returns, regardless of the content or sentiment of the news. For example, a news article with little new information might nevertheless make its information common knowledge. Such a news story could resolve information asymmetry and thereby change the liquidity of a market. Like the “dog that didn’t bark”, the mere fact that articles were published or not published about a firm contains information.

Another limitation of the decile spread results is that they do not reveal whether the predictability stems from positive or negative news. For example, Tetlock (2011) and Loughran and McDonald (2011) find that the preponderance of return response stems from negative news. In addition, to gauge the distinct effects of positive and negative news it is necessary to isolate the “news effect". In order to compare the potentially dissimilar effects of positive and negative news, we need to compare the effects to an appropriate return benchmark. The summary statistics in Table 2 show that firms without news have underperformed firms with news over our sample period. Inclusion of those firms using a portfolio methodology would bias the relative comparison of firms with different types of news. Specifically, firms with news would appear to outperform firms without news, regardless of the sentiment of the news, which would exaggerate the impact of positive news while reducing the apparent effect of negative news. In order to distinguish a publication effect from the quality of the information, and in order to separately evaluate the effect of positive and negative news, we must use a multivariate technique. This technique must separately measure the news effect, the effect of positive sentiment in news, and the effect of negative sentiment in news.

We use the cross-sectional regression technique of Fama and MacBeth (1973). For a given lag k ranging from 0 to 13, we regress stock returns on sentiment score

$$r_{i,t} = \alpha_{k,t} + \gamma_{k,t} * 1_{If_{news,t-k}} + \beta_{k,t} * Positive_{i,t-k} + \delta_{k,t} * Negative_{i,t-k} + \epsilon_{i,t} \quad (1)$$

where $r_{i,t}$ is the return on stock i in week t, $1_{If_{news,t-k}}$ is a dummy variable for firms with news over the given lag k, and $Positive_{i,t-k}$, $Neutral_{i,t-k}$ and $Negative_{i,t-k}$ are the evaluation of sentiment in news articles published in the lagged week. Following Fama (1976), we can interpret $\alpha_{k,t}$ as the return on an equally weighted portfolio of firms with no news at lag k. If "no news is good news," then $\alpha_{k,t}$ will tend to be negative. However, the summary statistics show that firms without news tend to underperform, so we expect this intercept to be positive. The term $\gamma_{k,t}$ represents the return premium for firms that have neutral published news over firms with no news. The $\beta_{k,t}$'s and $\delta_{k,t}$'s represent excess returns on costless, well-diversified portfolios that have 100% net loadings on positive or negative sentiment variables at a given lag.

===Insert Table 5 here====

Table 5 presents the average results of the regression coefficient time series, along with time series t-statistics using the Reuters sentiment engine. The "no news" intercept is negative at all lags, ranging from -1 basis point per week to -6 basis points per week. While these average returns are not statistically significantly different from zero, the consistently negative intercept shows that firms without news performed poorly over the sample period.

The premium for neutral news is $\gamma_k(t)$. It represents the weekly return premium of firms with 100% neutral news over firms with no news. The average point estimates are positive at all non-zero lags, showing there is a positive effect of neutral news. The positive coefficient for neutral news contradicts the well-known adage that "No News is Good News" popularized by Campbell and Hentschel (1992).

If neutral news is good, then we should expect positive news to be even better. The positive sentiment columns in Tables 5 confirm this intuition. The contemporaneous (lag 0) effect (β_0) is positive and highly statistically significant for all measures of sentiment. If news travels slowly, then good news should also have a positive lagged effect. However, this does not appear to be the case. The estimates for positive news in Table 5 are marginally significant at the 95% level at the first weekly lag but are not statistically positive at further lags. The subsequent point estimates are near zero and have different signs at higher lags. It appears that the market quickly incorporates positive information into returns.

Negative news also has a strong immediate effect on returns. In addition, there is a strong lagged effect shown in Table 5. The influence of Reuters sentiment is negative at all 13 lags, and the individual weeks are statistically significant at the 95% level at lags 1 through 6 and at lag 10. The pattern echoes the findings of Hong, Lim, and Stein (2000) that "bad news travels slowly". The findings are not consistent with Pound and Zeckhauser (1990) who find that rumor is already incorporated into prices.

===Insert Figure 5 here====

Figure 5 graphs cumulative coefficients from Table 5 for horizons ranging from 1 week to 13 weeks. The figure illustrates the pattern of news impact over different time periods. It shows that the effect of neutral news is small but accumulates positively for a full quarter. The incremental effect of positive sentiment is only positive for two or three weeks and then flattens out to negligible levels. In contrast, the impact of negative sentiment continues to be strong for the full 13 week period. Overall, neutral news, positive news, and negative news have different patterns of predicting stock returns through time. These findings demonstrate the importance of careful measurement of news sentiment and the distinct patterns of return predictability for positive, negative and neutral sentiment. The cross-sectional regression reinforces the portfolio results, which also show that a neural network predicts stock returns. The persistent predictive ability of negative Reuters sentiment is interesting in this regard. The findings are consistent with short sale constraints that prevent a small informed minority from fully impacting stock prices.

The previous section showed that abnormal returns persist when controlling for size and momentum. These variables are strongly correlated with the quantity and quality of news. Another relevant variable is earnings. In particular, Foster, Olsen, and Shevlin (1984) showed the existence of post-earnings-announcement drift, i.e., abnormally high returns in response to unexpectedly high earnings. The table of topic codes in the appendix shows that corporate financial results are the second most common topic tag (RES) in the database, and forecasts of financial results is the fourth most common topic (RESF). Even if news is not specifically about earnings, we can expect it to be correlated with earnings. If earning reports are a vehicle for quantifying and disseminating this news, then we might expect the good news to affect prices around the release of earnings, rather than before or after.

===Insert Table 6 here====

Table 6 addresses this issue by examining post-news returns relative to earnings announcement. In a week subsequent to a news story, we divide firms into three categories based on whether they have not yet announced earnings since the news ("Pre-earnings"), firms that announce earnings in that week ("Earnings"), and firms that have already announced earnings between the news release and the current week ("Post-Earnings"). Then we form decile spreads based on news sentiment within these three categories, as in Table 4. The Pre-earnings column of Table 6 shows that news sentiment does not predict statistically significant returns prior to firms' next earnings announcement. The cumulative abnormal return over 13 weeks in only 0.25%. To the extent that the Thomson Reuters Sentiment measures information, it appears that this information is not incorporated in stock prices prior to the next earnings release. But the Earnings column of Table 6 shows that quintile spread returns are mostly positive in earnings announcement weeks. Due to the smaller sample, the individual weeks are usually not statistically significant, but the cumulative excess return over 13 weeks is a healthy 5.57%, with a t-statistic of 2.7. This reinforces the results of Bernard and Thomas (1990) showing a delayed reaction to past earnings around subsequent announcements. The Post-Earnings quintile spreads are smaller, but still positive at 1.83% through Week 13, with a t-statistic of 2.1. This is consistent with the post-announcement earnings surprise anomaly of Foster, Olsen, and Shevlin (1984). Table 6 suggests that earnings announcements act as a channel of price discovery for information that was not immediately incorporated into stock prices when published. Hendershott, Livdan, and Schürhoff (2015) show that institutional trading anticipates news announcements. It is an open question whether institutions exploit news around the earnings announcements as well.

5 Conclusion

This paper investigates the usefulness of textual processing for predicting stock returns. We specifically use a neural network applied to a broad dataset of news stories. The duration of stock return predictability depends on the temporal aggregation of news. Predictability lasts only a few days when news is measured over day. But when we aggregate news over a week, the predictability lasts for up to a quarter year. The longer lasting predictability establishes that the effect of news on prices is not merely due to transient sentiment or liquidity. Instead, the deep textual analysis of the neural network appears to detect news that is persistently under-incorporated into current stock prices.

This paper distinguishes the effect of news from the positive or negative sentiment of that news. It also finds a news-attention effect, where firms with neutral news outperform firms without any news. Controlling for the news effect, this paper shows that positive news affects stock prices within one week. However, negative news predicts low stock returns for up to one quarter. This is consistent with short sale constraints that retard the incorporation of bad news. We find that most of the delayed reaction to news occurs around subsequent earnings announcements. This is consistent with earnings release and earnings-related trading acting as a channel to incorporate information into stock prices.

Future research can further explore patterns of predictability. For example, Chan (2003) found price reversals associated with returns that were unaccompanied by news, and Tetlock (2011) found overreaction to stale news. Boudoukh, Feldman, Kogan, and Richardson (2013) found differential response of returns to different types of news, including a greater response to relevant news. Other commercial products such as Ravenpack also analyze text. Comparison of return patterns across different types of news may enhance our understanding of how markets process information.

References

Akbas, F., E. Boehmer, B. Erturk, and S. M. Sorescu, 2013, "Short Interest, Returns, and Fundamentals," *SSRN Working Paper Series*, pp. 1–51.

Altman, E. I., G. Marco, and F. Varetto, 1994, "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)," *Journal of Banking & Finance*, 18(3), 505–529.

Antweiler, W., and M. Z. Frank, 2004, "Is all that talk just noise? The information content of internet stock message boards," *Journal of Finance*, 59(3), 1259–1294.

Atiya, A. F., 2001, "Bankruptcy prediction for credit risk using neural networks: A survey and new results," *IEEE Transactions on Neural Networks*, 12(4), 929–935.

Bernard, V. L., and J. K. Thomas, 1990, "Evidence that stock prices do not fully reflect the implications of current earnings for future earnings," *Journal of Accounting and Economics*, 13(4), 305–340.

Bollen, J., H. Mao, and X. Zeng, 2011, "Twitter mood predicts the stock market," *Journal of Computational Science*, 2(1), 1–8.

Boudoukh, J., R. Feldman, S. Kogan, and M. Richardson, 2013, "Which news moves stock prices? A textual analysis," *NBER Working Paper No. 18725*.

Brown, S. J., and J. B. Warner, 1985, "Using daily stock returns: The case of event studies," *Journal of financial economics*, 14(1), 3–31.

Butler, D., 2013, "When Google got flu wrong.," *Nature*, 494(7436), 155.

Campbell, J. Y., and L. Hentschel, 1992, "No news is good news: An asymmetric model of changing volatility in stock returns," *Journal of financial Economics*, 31(3), 281–318.

Chan, W., 2003, "Stock price reaction to news and no-news: Drift and reversal after headlines," *Journal of Financial Economics*, 70(2), 223–260.

Chen, H., P. De, J. Hu, and B.-H. Hwang, 2014, "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media," *Review of Financial Studies*, 27(5), 1367–1403.

Da, Z., J. Engelberg, and P. Gao, 2015, "The Sum of All FEARS Investor Sentiment and Asset Prices," *Review of Financial Studies*, 28(1), 1–32.

Das, S. R., and M. Y. Chen, 2007, "Yahoo! for Amazon: Sentiment extraction from small talk on the web," *Management Science*, 53(9), 1375–1388.

Dougal, C., J. Engelberg, D. García, and C. A. Parsons, 2012, "Journalists and the stock market," *Review of Financial Studies*, 25(3), 639–679.

Duhigg, C., 2006, "A smarter computer to pick stocks," http://www.nytimes.com/2006/11/24/business/24trading.html?pagewanted=print.

Dzielinski, M., and H. Hasseltoft, 2013, "Aggregate News Tone, Stock Returns, and Volatility," Unpublished Working Paper.

Fama, E. F., 1976, *Foundations of Finance: Portfolio Decisions and Securities Prices*. Basic Books.

Fama, E. F., and J. D. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81(3), 607–636.

Fang, L. H., and J. Peress, 2009, "Media coverage and the cross-section of stock returns," *Journal of Finance*, 64(5), 2023–2052.

Foster, G., C. Olsen, and T. Shevlin, 1984, "Earnings releases, anomalies, and the behavior of security returns," *Accounting Review*, pp. 574–603.

Garcia, D., 2013, "Sentiment during recessions," *Journal of Finance*, 68(3), 1267–1300.

Hagenau, M., M. Hauser, M. Liebmann, and D. Neumann, 2013, "Reading all the news at the same time: Predicting mid-term stock price developments based on news momentum," in *System Sciences (HICSS), 2013 46th Hawaii International Conference on*, pp. 1279–1288. IEEE.

Healy, A., and A. W. Lo, 2011, “Managing Real-Time Risks and Returns: The Thomson Reuters NewsScope Event Indices,” in *The Handbook of News Analytics in Finance*, ed. by G. Mitra, and M. L. John Wiley & Sons.

Hendershott, T., D. Livdan, and N. Schürhoff, 2015, “Are institutions informed about news?,” *Journal of Financial Economics*, fothcoming, 155.

Hillert, A., H. Jacobs, and S. Müller, 2014, “Media Makes Momentum,” *Review of Financial Studies*, 27(12), 3467–3501.

Hong, H., T. Lim, and J. Stein, 2000, “Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies,” *Journal of Finance*, 55(1), 265–295.

Infonic, 2008, “Reuters NewsScope Sentiment Engine,” Infonic Sentiment Technologies Whitepaper.

Jegadeesh, N., and S. Titman, 1993, “Returns to buying winners and selling losers: Implications for stock market efficiency,” *Journal of Finance*, 48, 65–91.

Jegadeesh, N., and D. Wu, 2013, “Word power: A new approach for content analysis,” *Journal of Financial Economics*, 110(3), 712–729.

Khandani, A. E., A. J. Kim, and A. W. Lo, 2010, “Consumer credit-risk models via machine-learning algorithms,” *Journal of Banking & Finance*, 34(11), 2767–2787.

Leinweber, D., and J. Sisk, 2011, “Event-Driven Trading and the "New News",” *Journal of Portfolio Management*, 38(1), 110–124.

Lerman, A., and J. Livnat, 2010, "The new form 8-K disclosures," *Review of Accounting Studies*, 15(4), 752–778.

Lo, A. W., 1994, "Neural Networks and Other Nonparametric Techniques in Economics and Finance," in *Blending Quantitative and Traditional Equity Analysis*, ed. by H. R. Fogler. Association for Investment Management and Research, chap. 2.

Loughran, T., and B. McDonald, 2011, "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance*, 66(1), 35–65.

Pound, J., and R. Zeckhauser, 1990, "Clearly heard on the street: The effect of takeover rumors on stock prices," *Journal of Business*, 63(3), 291–308.

Riordan, R., A. Storkenmaier, M. Wagener, and S. Sarah Zhang, 2013, "Public information arrival: Price discovery and liquidity in electronic limit order markets," *Journal of Banking & Finance*, 37(4), 1148–1159.

Ro, S., 2012, "We May Have Just Witnessed The Presence Of Artificial Intelligence In The Stock Market," http://www.businessinsider.com/art-cashin-articifial-intelligence-stock-market-2012-3.

Simon, D., and R. Heimer, 2012, "Facebook finance: How social interaction propagates active investing," in *AFA 2013 San Diego Meetings Paper*.

Sinha, N. R., 2016, "Underreaction to News in the US Stock Market," *Quarterly Journal of Finance*, 6(2), 1650005.

Tetlock, P., M. Saar-Tsechansky, and S. Macskassy, 2008, "More than words: quantifying language to measure firms' fundamentals," *Journal of Finance*, 63(3), 1437–1467.

Tetlock, P. C., 2007, "Giving content to investor sentiment: The role of media in the stock market," *Journal of Finance*, 62(3), 1139–1168.

——— , 2011, "All the News That's Fit to Reprint: Do Investors React to Stale Information?," *Review of Financial Studies*, 24(5), 1481–1512.

Uhl, M. W., M. Pedersen, and O. Malitius, 2015, "WhatŠs in the News? Using News Sentiment Momentum for Tactical Asset Allocation," *The Journal of Portfolio Management*, 41(2), 100–112.

von Beschwitz, B., D. B. Keim, and M. Massa, 2013, "Media-Driven High Frequency Trading: Evidence from News Analytics," *SSRN Working Paper Series*, pp. 1–46.

6 Tables and Figures

Table 1: Characteristics of News Sentiment Variables

Sentiment variable	Mean	Standard deviation
Thomson Reuters net sentiment	2.4%	39.0%
Thomson Reuters negative sentiment	27.5%	24.6%
Thomson Reuters positive sentiment	29.9%	21.7%

Notes: This table shows the average net firm sentiment (positive minus negative), positive sentiment and negative sentiment for 900,754 articles using the Thomson-Reuters sentiment engine.

Table 2: Weekly Summary Statistics by Market Capitalization

Decile	Log market cap	News stories per week	Proportion of firms without news	Return	Return w/news	Return w/o news	Difference	*t-statistics*
Smallest 1	9.42	0.16	0.95	-0.14%	2.00%	-0.24%	2.24%	3.47
2	10.56	0.59	0.94	-0.12%	1.51%	-0.23%	1.75%	3.59
3	11.22	1.33	0.92	-0.14%	-0.02%	-0.15%	0.12%	0.40
4	11.80	2.49	0.89	-0.06%	-0.04%	-0.06%	0.01%	0.04
5	12.35	4.01	0.87	-0.05%	-0.28%	-0.01%	-0.26%	-0.86
6	12.88	5.81	0.85	-0.01%	-0.26%	0.03%	-0.29%	-0.99
7	13.43	7.63	0.82	0.04%	-0.07%	0.06%	-0.13%	-0.48
8	14.04	10.19	0.76	0.05%	0.08%	0.04%	0.04%	0.15
9	14.83	13.30	0.66	0.09%	0.12%	0.07%	0.05%	0.25
Largest 10	16.48	22.42	0.34	0.06%	0.07%	0.04%	0.04%	0.18

Notes: This table presents weekly statistics for decile portfolios grouped on market capitalization over calendar years 2003 to 2010 (417 weeks). We divide the firms into deciles based on the market capitalization at the beginning of the month. Each week we note the number of news articles with relevance of at least 0.35, proportion of firms without news, average return, average return for firms with news, average return for firms without news, difference of return between firms with and without news and the *t-statistics* for the difference.

Table 3: Long-Short Excess Return from a Portfolio Based on News Sentiment on Day 0

Day after news	TR quintiles	
	Mean	t-statistics
-9	0.09%	6.0
-8	0.07%	4.5
-7	0.09%	5.9
-6	0.10%	6.3
-5	0.12%	7.4
-4	0.08%	4.7
-3	0.12%	6.9
-2	0.18%	10.8
-1	0.50%	22.4
0	1.99%	63.9
1	0.17%	9.8
2	0.04%	2.5
3	0.02%	1.2
4	0.04%	2.5
5	0.03%	1.6
6	0.06%	0.4
7	0.02%	1.1
8	0.01%	0.9
9	-0.02%	-1.2
10	-0.00%	0.0

Notes: We sort all stocks on a day based on the news sentiment from a lagged day and take a long position in the highest quintile (positive news stocks) and a short position in the lowest quintile (negative news stocks). This table shows the average daily return and t-statistics on long-short portfolio from sentiment scores using the Thomson-Reuters sentiment engine.

Table 4: Weekly Returns from Long-Short Portfolio Based on News in Week 0.

(a) Panel A: Long-short excess returns from weekly portfolio for all stocks with news

Week after news	Return	t-statistics	Momentum-adj return	t-statistics	Size-adj return	t-statistics
0	3.75%	37.0	3.61%	41.9	3.62%	36.8
1	0.32%	3.9	0.31%	2.5	0.36%	2.5
2	0.20%	2.6	0.12%	1	0.20%	1.4
3	0.26%	3.6	0.22%	1.7	0.25%	1.8
4	0.10%	1.4	0.01%	0.1	0.02%	0.1
5	0.19%	2.6	0.13%	1	0.21%	1.5
6	0.14%	1.9	0.22%	1.8	0.25%	1.9
7	0.11%	1.5	0.00%	0	0.02%	0.2
8	0.08%	1.2	0.14%	1.1	0.19%	1.4
9	0.12%	1.6	0.23%	1.8	0.24%	1.8
10	0.21%	2.8	0.23%	1.6	0.22%	1.5
11	0.20%	2.9	0.29%	2.3	0.36%	2.6
12	0.01%	0.2	0.05%	0.4	0.06%	0.5
13	0.21%	2.6	0.27%	2.1	0.27%	2.0

(b) Panel B: Excess weekly returns from long-short portfolio by days of news in week 0.

Week after news	One article	t-statistics	Multiple articles	t-statistics
0	3.37%	41.08	4.22%	24.46
1	0.19%	3.13	0.48%	3.94
2	0.11%	1.98	0.21%	1.79
3	0.15%	2.92	0.21%	1.86
4	0.08%	1.29	-0.04%	-0.36
5	0.09%	1.67	0.20%	1.85
6	0.09%	1.56	0.10%	1.00
7	0.12%	2.09	0.01%	0.14
8	0.06%	1.13	0.13%	1.24
9	0.08%	1.33	0.21%	1.88
10	0.11%	2.05	0.13%	1.17
11	0.11%	2.06	0.06%	0.64
12	-0.05%	-0.99	0.07%	0.69
13	0.10%	1.69	0.25%	2.32

Notes: We sort all stocks in a week based on the news sentiment from Week 0 and take a long position in the highest decile (positive news stocks) and a short position in the lowest decile (negative news stocks). Panel A shows the average weekly return on long-short portfolio using sentiment scores using the Thomson Reuters (Reuters) sentiment engine as well as returns adjusted for 26-week momentum and logarithm of market capitalization. To control for momentum, we assign stocks to ten momentum deciles based on returns over past 26 weeks, and calculate the benchmark return for each momentum decile. For each stock, we then calculate the excess return over its benchmark. Long-short returns are reported in excess of the benchmark return. We similarly adjust for size. In Panel B, the One Article column shows the average excess return from a long-short portfolio of stocks that had only one news article in week 0. The Multiple Articles column shows the excess return from a long-short portfolio of stocks with more than one news article in week 0.

Table 5: Cross-Sectional Regressions of Weekly Returns based on Thomson Reuters Sentiment in Week 0

Average cross-sectional regression coefficient on Thomson Reuters sentiment variables

Week after news	α	t-statistics	News effect	t-statistics	Positive sentiment	t-statistics	Negative sentiment	t-statistics
0	-0.0005	-0.3	-0.0003	-0.4	0.0304	27.6	-0.0329	-23.5
1	-0.0005	-0.3	0.0008	1.4	0.0017	2.0	-0.0037	-3.7
2	-0.0006	-0.4	0.0008	1.5	0.0008	0.9	-0.0018	-1.9
3	-0.0006	-0.4	0.0012	2.2	0.0004	0.5	-0.0032	-3.5
4	-0.0005	-0.3	0.0019	3.4	-0.0016	-1.9	-0.0027	-2.7
5	-0.0004	-0.3	0.0012	2.2	-0.0003	-0.4	-0.0026	-2.9
6	-0.0003	-0.2	0.0009	1.5	-0.0002	-0.2	-0.0018	-1.9
7	-0.0003	-0.2	0.0010	1.8	0.0002	0.2	-0.0015	-1.6
8	-0.0003	-0.2	0.0012	2.1	-0.0005	-0.7	-0.0018	-1.9
9	-0.0002	-0.2	0.0004	0.7	0.0007	0.9	-0.0011	-1.1
10	-0.0001	0.0	0.0011	2.0	-0.0005	-0.7	-0.0026	-2.8
11	-0.0004	-0.3	0.0003	0.6	0.0009	1.2	-0.0011	-1.3
12	-0.0004	-0.3	0.0013	2.4	-0.0005	-0.7	-0.0009	-1.0
13	-0.0002	-0.2	0.0004	0.8	0.0008	1.0	-0.0017	-1.8
					0.0321		-0.0593	

Notes: We use the cross-sectional regression technique of Fama and MacBeth (1973). For a given lag k ranging from 0 to 13, we regress stock returns on Thomson Reuters sentiment scores

$$r_{i,t} = \alpha_{k,t} + \gamma_{k,t} * 1_{If_{news,t-k}} + \beta_{k,t} * Positive_{i,t-k} + \delta_{k,t} * Negative_{i,t-k} + \epsilon_{i,t} \quad (2)$$

where $r_{i,t}$ is the return on stock i in week t, $1_{If_{news,t-k}}$ is a dummy variable for firms with news over the lag k, and $Positive_{i,t-k}$, and $Negative_{i,t-k}$ are the evaluation of positive and negative sentiment in news article published in the lagged week k. We report the time series average of $\gamma_{k,t}$ in the 'News Effect' column and report the time series averages of $\beta_{k,t}$ and $\delta_{k,t}$ in the 'Positive Sentiment' and 'Negative Sentiment' columns..

Table 6: Long-Short Excess Return Pre-earnings, During Earnings and Post-earnings Weeks

Week after news	Pre-Earnings		Earnings week		Post-Earnings	
	Average	t-statistics	Average	t-statistics	Average	t-statistics
0	3.40%	31.34	5.90%	13.39		
1	0.32%	1.88	0.81%	1.71		
2	0.09%	0.42	0.80%	1.27	0.04%	0.08
3	0.30%	1.32	1.44%	2.91	0.37%	1.05
4	0.10%	0.35	-0.13%	-0.21	-0.31%	-1.19
5	0.31%	1.01	0.57%	0.81	-0.08%	-0.39
6	0.07%	0.2	1.16%	1.82	0.23%	1.12
7	-0.89%	-2.21	-0.20%	-0.37	0.04%	0.21
8	-0.23%	-0.51	0.09%	0.17	0.13%	0.79
9	0.24%	0.44	0.66%	0.93	0.15%	0.9
10	-0.19%	-0.27	0.10%	0.18	0.23%	1.27
11	-0.95%	-1.17	0.93%	1.75	0.27%	1.77
12	0.00%	0	-1.52%	-2.91	0.42%	2.67
13	1.09%	0.81	0.85%	1.73	0.36%	2.26
Week 1-13	0.25%	0.11	5.57%	2.66	1.83%	2.14

Notes: We sort all stocks in a week based on the news sentiment from a lagged week and take a long position in the highest decile (positive news stocks) and a short position in the lowest decile (negative news stocks). Pre-earnings column shows the average weekly return on long-short portfolio using sentiment scores before the latest earnings release. Earnings column shows the long-short return from a weekly long-short strategy during the earnings week. Post-earnings column shows the return after the latest earnings news. The last row shows the average return for news long-short strategy pre-earnings, during earnings and post-earnings period.

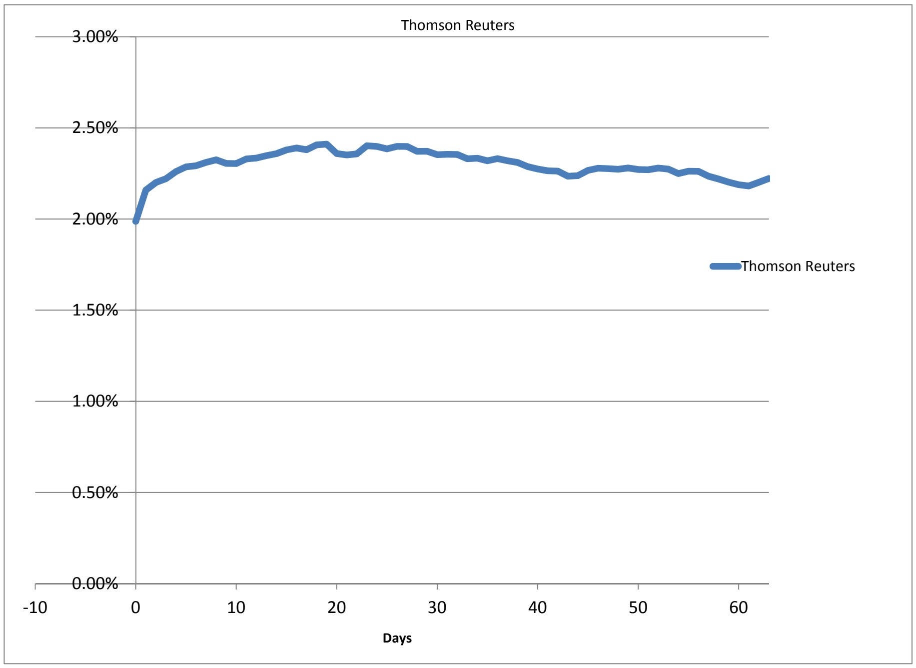


Figure 1: Cumulative Daily News and Post-News Long-Short Quintile Returns

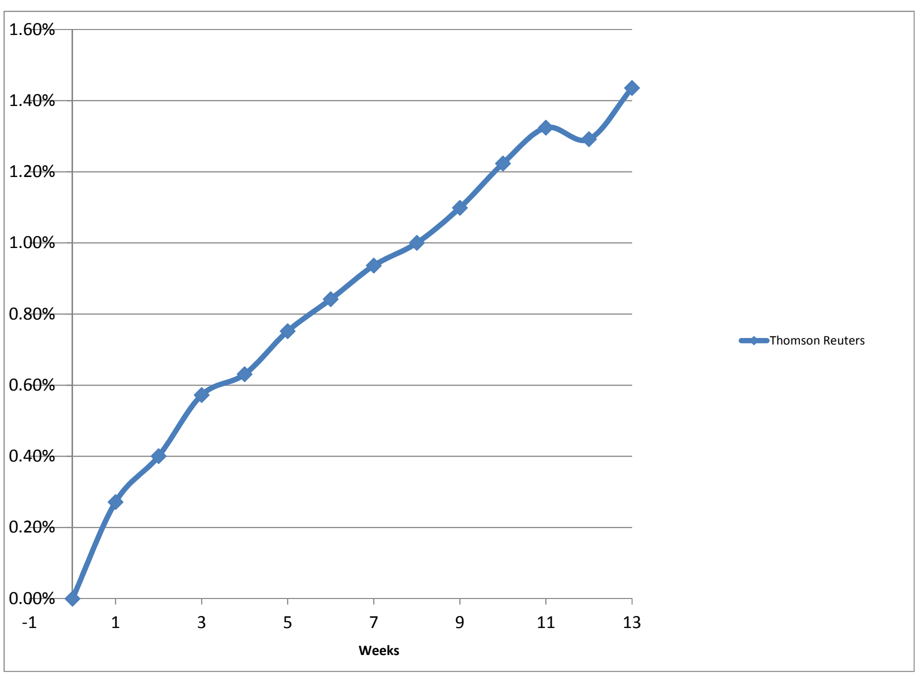
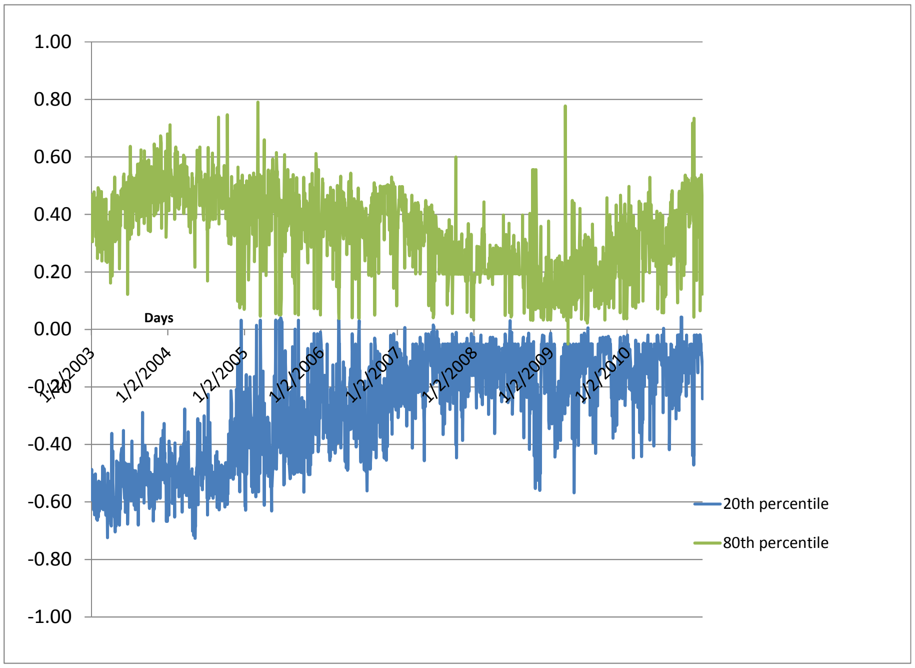
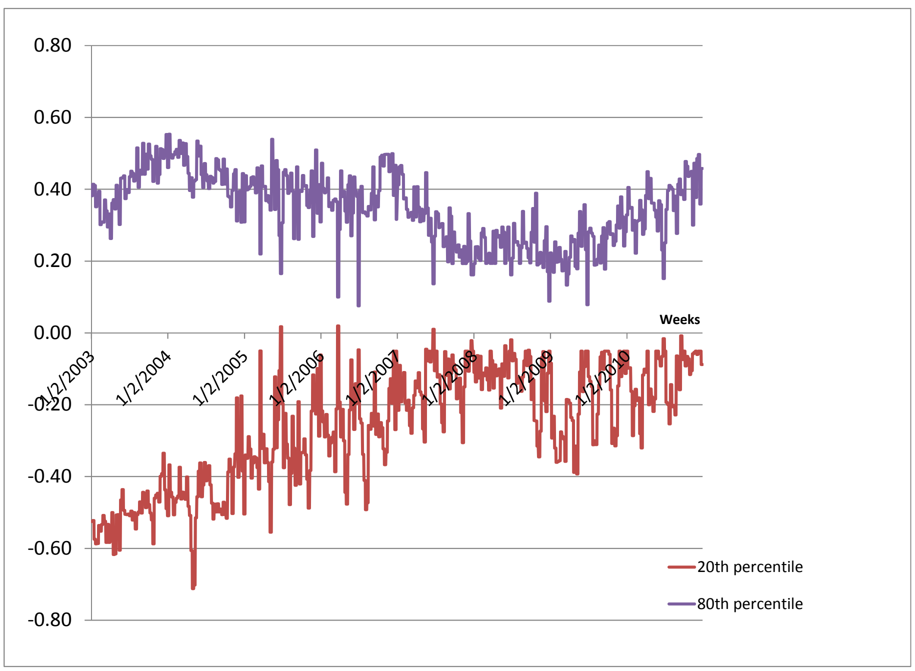


Figure 2: Cumulative Weekly Post-News Long-Short Decile Returns



(a) Daily Fractile Levels for Thomson-Reuters News Sentiment



(b) Weekly Fractile Levels for Thomson-Reuters News Sentiment

Figure 3: Daily and Weekly Fractile Levels for Thomson-Reuters News Sentiment

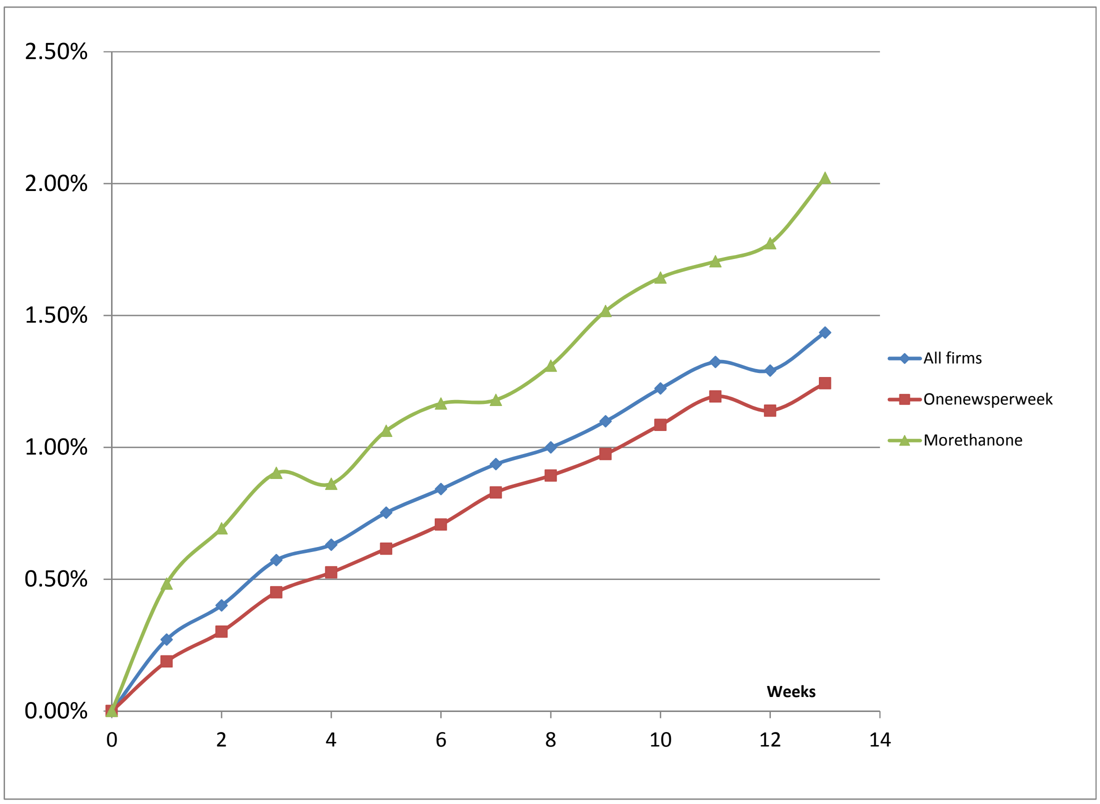
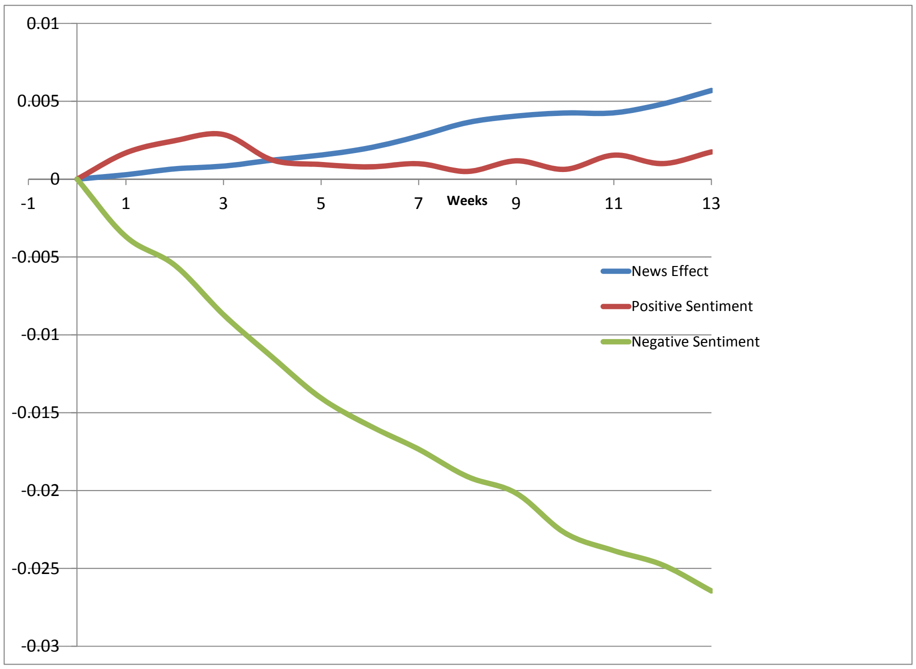


Figure 4: Cumulative Weekly Post-News Long-Short Quintile Returns

Figure 5: Cumulative Weekly Average Cross-Sectional Regression Coefficients



The figure plots the cumulative coefficients from Table 5 for horizons ranging from one week to thirteen weeks. The table reports time series average from the following regression. For a given lag k ranging from 0 to 13, we regress stock returns on sentiment ratings

$$r_i(t) = \alpha_k(t) + \gamma_k(t) * 1_{If_{news}}(t-k) + \beta_k(t) * Positive_i(t-k) + \delta_k(t) * Negative_i(t-k) + \epsilon_i(t)$$

where $r_i(t)$ is the return on stock i in week t, $1_{If_{news}}(t-k)$ is a dummy variable for firms with news over the lag k, and $Positive_i(t-k)$, and $Negative_i(t-k)$ are the evaluation of positive and negative sentiment in news article published in the lagged week k.

Table A1: List of Topic Codes

Topic Code	Brief Description	Percentage of news
MRG	Mergers and Acquisitions (including changes of ownership)	17.7%
RES	Corporate Results	14.0%
STX	Stock Markets	13.4%
RESF	Corporate Results Forecasts	9.2%
NEWS	Major breaking news	8.2%
DBT	Debt Markets	6.3%
RCH	Broker Research and Recommendations	4.9%
HOT	Hot Stocks	4.9%
CORA	Corporate Analysis	4.4%
INV	Investing	3.0%
REGS	Regulatory Issues	2.7%
PRO	Biographies, Personalities, People	2.1%
MNGISS	Management issues/policy	1.6%
AAA	Ratings	1.5%
DIV	Dividends	1.4%
PRESS	Press Digests	1.4%
IPO	Initial Public Offerings	1.0%
WIN	Reuters Exclusive News	0.7%
ECI	Economic Indicators	0.5%
EXCA	Exchange Activities	0.5%
BKRT	Bankruptcies	0.3%
RSUM	Reuters Summits	0.2%
FED	Federal Reserve Board	0.2%
CFIN	Corporate Finance	0.0%
ERR	Error	0.0%
FES	Editorial Special, Analysis and Future Stories	0.0%
INSI	Technical Analysis	0.0%
TRN	Translated News	0.0%
CONV	Convertible Bonds	0.0%
CDM	Credit Market News	0.0%
NEWR	Original Corporate News Releases	0.0%
DDEAL	Directors' Dealings	0.0%
DIARY	Diaries	0.0%
KEY	Key Personnel Moves at Corporations or Banks	0.0%
TOP	Top News	0.0%

Table provides a listing of all "Topic Codes" in our news sample. Column 1 provides the topic code tag, column 2 provides a brief description, column 3 provides the percentage of news item with that tag. Many news items have more than one news tag.