The Momentum of News

Ying Wang, Bohui Zhang, and Xiaoneng Zhu*

This Draft: December 2016

*Ying Wang is from School of Finance, Central University of Finance and Economics, 36 South College Road, Beijing, China, 100086; Bohui Zhang is from the School of Banking and Finance, UNSW Business School, UNSW Australia, Sydney, NSW, Australia, 2052, and Xiaoneng Zhu is from School of Finance, Shanghai University of Finance and Economics, Guoding Road 777, Shanghai, China, 200433. Authors' contact information: Wang: wang0371@e.ntu.edu.sg, (86)10 62288607; Zhang: bohui.zhang@unsw.edu.au, (61) 2-93855834; Xiaoneng Zhu: xiaonengz@gmail.com, (86) 21 65908162. We are grateful for the valuable comments from Jie Cao, Liyan Han, Andrew Karolyi, Xindan Li, Bo Liu, Neng Wang, Yu Yuan, and the conference and seminar participants at the 2016 Conference on Corporate Finance and Capital Markets in Shanghai and the 2016 Chinese Finance Annual Meeting in Dalian. We also thank the program committee of the 2016 Chinese Finance Annual Meeting for awarding us the Best Paper Award. Bohui Zhang acknowledges the research grants from the ARC discovery grant (DP 120104755) and ARC linkage grant (LP130101050) from the Australian Research Council and the CIFR research grants (E026 and E028) from the Centre for International Finance and Regulation. Xiaoneng Zhu acknowledges the financial support from the Natural Science Foundation of China (Grant No. 71473281).

The Momentum of News

Abstract

Relying on a comprehensive data set of news releases, we construct monthly firm-level news sentiment scores during the 2000–2014 period and document a news momentum phenomenon that stocks with more positive news in the past generate more positive news in the future. We propose two hypotheses to explain this phenomenon and find that news momentum is driven by the persistence of firms' fundamentals instead of firms' information environments. A trading strategy, which combines a long position in a good-news quintile portfolio with a short position in a bad-news portfolio, generates 8.352 percent risk-adjusted return annually. This return anomaly appears on both news and non-news days. Overall, these findings suggest that the cross-sectional prediction of news is not fully incorporated into the stock price by investors.

Keywords: News; Momentum; Fundamentals; Information Environments; Future Returns
JEL Code: G02; G10; G14

1. Introduction

Over the past four decades, hundreds of anomalies have been uncovered in the cross-section of stock returns. Among potential explanations for cross-sectional predictability, mispricing is identified as the key one (e.g., Mclean and Pontiff, 2016; Engelberg, McLean, and Pontiff, 2016). In particular, behavioral theories attribute mispricing to investors' inability to price news correctly (e.g., Daniel, Hirshleifer, and Subrahmanyam, 1998; Barberis, Shleifer, and Vishny, 1998; Hong and Stein, 1999). Because these theories typically take news as given, the property of news is left unexplored. Given price movement is a function of news, the predictability of news is essential to the understanding of return anomalies. In this paper we fill this void by examining the cross-sectional predictability of news.

Using a comprehensive news dataset collected by RavenPack, we construct a sample of real-time news releases for stocks listed on the NYSE, Amex, and Nasdaq over the 15-year period between 2000 and 2014. We focus on news articles commonly used by institutional and sophisticated individual investors. Specifically, RavenPack quantifies the positive (or negative) information (i.e., news sentiment score) in each news article based on professional algorithms. For example, a news article on a corruption scandal involving a firm's executives is associated with a low news sentiment score, and a news article regarding the successful development of a firm's new product is associated with a high news sentiment score. Our main analysis is conducted at the monthly frequency. We aggregate news articles for each firm in a month and then calculate monthly news sentiment scores by averaging news sentiment scores over a month.

We perform the following analyses. First, we examine whether there is a cross-sectional pattern of news. Specifically, we construct monthly news portfolios by sorting stocks into five quintile portfolios based on their current news sentiment scores. We then compute the equally-weighted average news scores of each portfolio. We find that stocks in the highest news sentiment score portfolio outperform stocks in the lowest news sentiment score portfolio in terms of their news sentiment scores in the future, which is called *the news momentum phenomenon*. This phenomenon is robust to various specifications such as the daily or weekly frequency, the inclusion of neutral news articles, negative or positive news sorting, and decile portfolios.

Second, we propose two hypotheses to explain the news momentum phenomenon. One view is that news momentum is driven by firms' information environments. For example, stale news is likely to be disseminated over and over again. The release of stale information is meaningful when some investor cannot appropriately distinguish between new and old information (e.g., Tetlock, 2011) due to limited attention or other reasons. Moreover, when insiders' disclosure preferences are not aligned with those of outside investors, various incentives can motivate managers to strategically disclose or withhold firm-specific information (e.g., Kothari, Shu, and Wysocki, 2009). These properties of news dissemination associated with information environments may induce news momentum that companies with more positive information continue to disseminate more positive news, and companies with less positive information continue to disclose less positive news. We call this view *the information environment hypothesis*.

One the other hand, news momentum could be induced by the dynamics of firms' fundamentals. It has been well documented that earnings are persistent and predictable (e.g., Ball and Brown, 1968; Beaver, Clarke and Wright, 1979; Fama and French, 2000, 2006; Markov and Tamayo, 2006; Li, 2010; Novy-Marx, 2015). If news articles fairly reflect firms' fundamentals, the persistent earning stream is likely to generate a persistent stream in sequential news releases. More specifically, positive (negative) news is more likely to be followed by positive (negative) fundamentals, which is delivered in news. We call this view *the fundamental hypothesis*.

To test the two hypotheses, we start by checking whether news momentum concentrates among stocks with poor information environments proxied by small firm size, low analyst coverage, and less institutional holdings. Inconsistent with *the information environment hypothesis*, we find no systematic difference in news momentum between stocks with small firm size, low analyst coverage, and less institutional holdings and stocks with large firm size, high analyst coverage, and more institutional holdings. We move forward by testing whether news momentum is driven by firms' fundamentals. In supportive of *the fundamental hypothesis*, we find that firms with current good news scores have higher profitability or earnings surprises in the future. When we further decompose news into hard news (news is more relevant to firms' fundamentals) and soft news (news is less relevant to firms' fundamentals), we show that the fundamental prediction is mainly driven by hard news.

Finally, we investigate whether investors are aware of news momentum. If news momentum is correctly incorporated into the stock price, stocks in the highest news score portfolio should have similar future returns as stocks in the lowest news score portfolio. Interestingly, we find significant news-driven return predictability: the strategy that buys the good news portfolio and sells bad news portfolio generates a return of 8.35 percent per year. News-driven return predictability is only significant in short horizons and in stocks with poor information environments such as stocks with small firm size, low analyst coverage, and less institutional holdings. This is consistent with the mispricing view of return predictability.

In the context of mispricing, there are two forms of underreaction that can account for observed return predictability. The first form of mispricing is that market participants do not realize the presence of news continuation and underestimate the persistence of news (fundamentals). The other form is that investors underreact to current news and induce return continuation such as the post-earnings announcement drift anomaly. To test two types of investor underreaction, we decompose future returns into returns on news days and non-news days. We show that news-driven return predictability is caused by underreaction to both news momentum and news itself.

Finally, we provide additional tests and show that the return predication of news is more pronounced for hard news than soft news and is robust to various specifications such as the Fama and MacBeth (1973) approach, the daily or weekly frequency, the inclusion of neutral news articles, negative or positive news sorting, and decile portfolios.

Our paper contributes to two strands of the literature. First, our paper is related to the literature on the capital market impact of business media. The media disseminates or rebroadcasts a large amount of financial news or signals regarding to firms' earnings, management, and investment decisions, among others. These pieces of information affect investors' expectations about stock returns and may improve market efficiency. Indeed, a flood of research highlights the informational role of the media through various channels such as drawing attention (Fang and Peress, 2009; Da, Engelberg, and Gao, 2011), resolving information asymmetry (e.g., Tetlock, 2010), delivering fundamental information (e.g., Tetlock, Saar-Tsechansky, and Macskassy, 2008), or inflating market sentiment (e.g., Tetlock, 2007). Our study documents a cross-sectional pattern of news and provides an asset pricing implication of the news pattern.

Second, our paper contributes to the literature on the predictability of stock returns based on news stories. Hillert, Jacobs, and Muller (2014) show that firms covered by the media exhibit stronger return momentum, indicating that news dissemination exacerbates investor biases. Using a comprehensive sample of intraday firm-specific news data, Jiang, Li, and Wang (2015) decompose stock returns to news-driven and non-news-driven components. They find that the news-driven return is particularly pronounced for firms with less analyst coverage, higher volatility, and lower liquidity. This is consistent with imperfect investor reaction to news and limits to arbitrage. More broadly, using a sample of 97 stock return anomalies documented in published studies, Engelberg, McLean, and Pontiff (2016) show that anomaly returns are seven times higher on earnings announcement days and two times higher on corporate news days. We offer a new insight that investors do not only misprice news but also misprice the cross-sectional pattern of news.

The remainder of the paper proceeds as follows. We explain the sample construction for the news variable and describe sample characteristics in Section 2. In Section 3, we examine news momentum and test two hypotheses on news momentum. In Section 4, we study the return prediction of news momentum. Finally, we provide concluding remarks in Section 5.

2. Data, variable construction, and descriptive statistics

2.1. Data and sample

Our data come from a variety of sources. First, we obtain stock returns and market capitalization data from CRSP and firm fundamentals and earnings announcement data from Compustat. After merging the CRSP and Compustat data, our initial sample covers all common stocks listed on the New York Stock Exchange (NYSE), American Stock Exchange (Amex), and National Association of Securities Dealers Automated Quotations (Nasdaq). We further require stocks to have the stock price greater than one dollar and non-missing information on market capitalization and the book-to-market ratio.

Then we obtain news data from RavenPack News Analytics, and include only firms with at least one news story covered by RavenPack. RavenPack is a leading global news database used by practitioners in quantitative and algorithmic trading and also by scholars in accounting and finance research (e.g., Kelley and Tetlock, 2013; Kolasinski, Reed, and Ringgenberg, 2013; Schroff, Verdi, and Yu, 2014; Dai, Parwada, and Zhang, 2015, Dang, Moshrian, and Zhang, 2015; Jiang, Li, and Wang, 2015). RavenPack collects and analyzes real-time, firm-level business news from leading news providers (e.g., *Dow Jones Newswire*, *The Wall Street Journal*, and *Barron's*) and other major publishers and web aggregators, including industry and business publications, regional and local newspapers, government and regulatory updates, and trustworthy financial websites.

Our final sample consists of 473,941 firm-month observations (2,662 firms on average) spanning from January 2000 to October 2014. The sample period is determined by the availability of RavenPack. In some of additional analyses, we also use a battery of firm characteristics including analyst coverage, institutional ownership, the return of asset, earnings surprises, idiosyncratic volatility, beta, past stock returns, and illiquidity (Amihud, 2002). These firm characteristics either proxy for firms' information environments or predict stock future returns. Appendix A lists data sources for these characteristic variables.

2.2. News sentiment

Our analysis is conducted based on the enormous volume of news flows: 5,610,988 news articles in total, which is equivalent to about 31,522 news articles per month or 12 news articles per firm and month. To measure the informational content of a news article, RavenPack implements two steps. First, RavenPack classifies the news articles into news event categories according to the RavenPack taxonomy. Both the topic and a firm's role in the news article are tagged and categorized. For example, a news article with the headline "IBM Completes Acquisition of Telelogic AB" is categorized into the "acquisition" event, and tagged as "acquisition-acquirer" for IBM and as "acquisition-acquiree" for Telelogic AB. Similarly, a news article titled "Xerox Sues Google over Search-Query Patents" is categorized into the "patent-infringement". Xerox has the tag "patent-infringement-plaintiff" while Google obtains the tag "patent-infringement-defendant". All the news articles in our study can be grouped into 32 news categories, which are listed in Appendix B. Among all news stories, the top five news categories are "earnings" (19.55%), "insider-trading" (14.07%), "products-services" (8.47%), "technical-analysis" (7.42%), and "order-imbalances" (6.99%).

Second, RavenPack constructs the news sentiment score for each news article based on professional algorithms, which were developed and evaluated by effectively combining traditional language analysis, financial expert consensus, and market response methodologies.[1] Specifically, the news sentiment score indicates whether or not, and to what extent a news story may have a positive, neutral, or negative effect on stock prices. This score is assigned to all relevant firms listed in the news report. The sentiment score ranges from 0 to 100, with a value below (above) 50 indicating the negative (positive) sentiment of a given news. A score of 50 represents a neutral sentiment. To facilitate our empirical analysis, we minus 50 from the news sentiment scores and scale it by 50. After the adjustment, the adjusted sentiment scores fall in an interval between -1 and 1.

The news sentiment score is economically intuitive. For example, a firm has a news sentiment score of -0.2 for a news article about its analyst downgrade from “Buy to a Neutral”. In terms of relative magnitude, the firm may obtain a more negative news sentiment score such as -0.4 if the analyst downgrade is from “Strong Buy to a Strong Sell”. For complicated news stories including financial variables and economic indicators, the news sentiment score can intellectually measure the percentage change between announced actual figures and the market consensus (or any other benchmarks). For example, a firm beating earnings by 70% may enjoy a news sentiment score of 0.6 while the firm exceeding a benchmark by 1% may have a news sentiment score of 0.1.

We exclude repeated news by setting the event novelty score (ENS) provided by RavenPack to be 100, which captures only fresh news about a firm. As such, the cross-sectional pattern of news is unlikely to be driven by reproduction or re-dissemination of the same or similar articles. Our main analysis is conducted at the monthly frequency. To obtain monthly observations, we calculate the sum of news sentiment scores for each firm over all the news events during the month and then use this number scaled by the number of news as the key measure of news sentiment (*News*). The news score is set to be zero if there is no news for a firm in a particular month.

To differentiate news articles with respect to their relevance to firms’ fundamentals, we split news articles into two groups: hard news and soft news. The hard news group consists of four

[1] For example, RavenPack algorithms can analyse actual figures, estimates, ratings, revisions, magnitudes, and recommendations disclosed in news articles. Its algorithms can also compare actual with estimated figures about earnings, revenues or dividends and produce the news sentiment score based on these comparisons. In addition, these algorithms can calculate percentage differences between financial figures and analyse stock and credit ratings disclosed by analysts. These algorithms can also process information such as the Richter scale in the case of an earthquake or the number of casualties in a suicide bombing event. The use of emotionally charged language by authors is also incorporated into these algorithms when shaping the strength component of the news sentiment score.

news categories: “revenues”, “earnings”, “analyst-ratings”, and “credit-ratings”. All other news categories are included in the soft news group. Appendix B shows the details of two news groups. Among all news articles, 30.4% are defined as hard news and the rest of 69.6% are soft news. Following the same procedure of *News*, we calculate the sentiment of hard news (*HardNews*) and the sentiment of soft news (*SoftNews*) for each firm and month.

2.3. Summary statistics

Table 1 reports the descriptive statistics of the main variables used in our empirical analysis. In our sample, we exclude firms with zero news scores each month.[2] We find that, on average, firms in our sample have positive news ($News_t$=0.083). Another notable observation that emerges from the table is the symmetric distribution of our news sentiment variable. For example, the mean of $News_t$ is equal to the median of $News_t$. Moreover, hard news and soft news have a similar distribution. For example, the means of $HardNews_t$ and $SoftNews_t$ are 0.109 and 0.080, respectively; the standard deviations of $HardNews_t$ and $SoftNews_t$ are 0.260 and 0.151, respectively.

In addition to these news variables, the table also shows reasonable summary statistics for other variables: next month returns (R_{t+1}), logarithm of market capitalization ($Size_t$), analyst coverage ($Analyst_t$), institutional ownership ($InstOwn_t$), book-to-market ratio (B/M_t), beta ($Beta_t$), idiosyncratic volatility ($IdioVol_t$), past two-month stock returns ($Return_{t-3,\ t-2}$), past three-month stock returns ($Return_{t-6,\ t-4}$), past six-month stock returns ($Return_{t-12,\ t-7}$), Amihud’s (2002) illiquidity ($Illiquidity_t$), earnings surprise (SUE_t), and ROA (ROA_t).

[Insert Table 1 Here]

Table 2 provides an overview of the extent of news coverage for five periods (2000-2002, 2003-2005, 2006-2008, 2009-2011, and 2012-2014) across size portfolios. Firms are sorted into size quintiles by their market capitalization. Panel A reports the number of news articles covering a specific firm in a month. It is evident from the panel that large firms have a larger number of

[2] About 25.8% of total firm-month observations have zero news scores, indicating that those firms have either neutral news or no news in the month. However, our main results always hold even including those zero scores observations in the analysis, which is shown in our robustness tests (see Internet Appendix Table IA2 and Table IA3),

news articles. For example, in the 2012-2014 period, firms in the large *Size* quintile have approximately 36 news articles in a month, while firms in the small *Size* quintile only have eight news articles per month. This fact is not only consistent with the media literature that large firms attract higher media coverage, but also aligned with the intuition that large firms typically generate more news events.

We also find that the number of news articles increases substantially over the past 15 years across all size portfolios. For example, the number of news articles covering firms in the large *Size* quintile increases from nine in the 2000-2002 period to 36 in the 2012-2014 period. This time-series pattern could be explained by either the increasing intensity of media coverage or growing firm activities.

Panel B and C present the number of positive and negative news articles, respectively. Overall, we find similar cross-sectional and time-series patterns for both positive and negative news articles as Panel A.

[Insert Table 2 Here]

3. The cross-section of news

3.1. Baseline results

This section investigates the cross-sectional pattern of news. Following the standard method in the asset pricing literature, we construct monthly news portfolios according to firms' news sentiment scores. Specifically, at the end of month *t*, we sort all stocks into five portfolios based on their news sentiment scores in month *t* ($News_t$). We then compute the equally-weighted average news scores for the current month *t* and future months from *t*+1 to *t*+24 across all firms in each portfolio. The quintile of stocks releasing the most negative information (below the 20th percentile) is named the bad *News* portfolio, and the quintile of stocks releasing the most positive information (above the 80th percentile) is called the good *News* portfolio.

Table 3 presents the cross-section of news. The bad *News* portfolio has a sentiment score of -0.134 at the formation period, while the good *News* portfolio has a corresponding sentiment score of 0.296. The 2nd to 4th quantiles of stocks have respectively news sentiment scores of 0.009, 0.083, and 0.159. To encourage the comparison between the bad *News* portfolio and the good *News* portfolio, we construct a hedging portfolio by selling stocks in the bad *News* portfolio and buying stocks in the good *news* portfolio. By construction, the “good minus bad” (GMB) *News* portfolio has a positive sentiment score of 0.431, which is statistically significant at the 1% level.

[Insert Table 3 Here]

How does the sentiment of the GMB portfolio fluctuate? If we hold the GMB *News* portfolio for a number of months from *t*+1 to *t*+24 following the formation month, the GMB *News* portfolio exhibits a pattern of news sentiment continuation in the subsequent periods. We first look at the GMB portfolio sentiment at time *t*+1. The GMB *News* portfolio has an average positive sentiment score of 0.037, with a heteroscedasticity and autocorrelation consistent *t*-value of 25.7. The magnitude of news sentiment is economically significant, which is equivalent to the 23.6% of the standard deviation of *News*. Indeed, this pattern of news sentiment continuation not only presents in the GMB portfolio, any other hedging portfolios constructed by buying some basic news portfolio and selling another news portfolio exhibits such a pattern. Momentum, the tendency of an object in motion to stay in motion, accurately describes this pattern. We hence label this cross-sectional pattern of news sentiment movement *the momentum of news* or *news momentum*.

Turning to the holding period from *t*+2 to *t*+6, which removes the impact of news sentiment at period *t*+1, we find that the GMB portfolio still exhibits a positive sentiment score of 0.044 with a *t*-statistic of 49.8, suggesting that news sentiment stays in motion. Furthermore, all quantile portfolios also show a monotonic relation in terms of news sentiment scores. Looking at the holding period from *t*+7 to *t*+12 and from *t*+13 to *t*+24, it is evident that the GMB *news* portfolio has a significant positive sentiment score.[3] For these holding periods, we also find a monotonic relation in term of sentiment among all quantile portfolios, though news sentiment scores of these portfolios converge more quickly.

Overall, our empirical findings suggest the presence of news momentum,[4] which implies that firms with more positive (negative) news are likely to release more positive (negative) news in the future. This phenomenon is likely to have important implications for stock prices, it is thus

[3] At longer horizons, news momentum still continues for some periods and finally becomes insignificant.
[4] We examine the robustness of the momentum of news by using different specifications in the Internet Appendix Table IA1.

important to understand the economic driving force of the news momentum.

3.2. Hypothesis development

This section develops two hypotheses regarding the economic mechanisms of news momentum formation. Our first conjecture is that news momentum is driven by firms' information environments. For example, the transmission of stale news may induce news momentum. In a competitive information market, news is a non-rival good with a high discovery cost and a low replication cost (e.g., Veldkamp, 2006). This makes the dissemination of old news very cheap. As such, stale news is likely to be disseminated over and over again. In particular, the release of stale information is meaningful when some investor cannot appropriately distinguish between new and old information (e.g., Tetlock, 2011) due to limited attention or other reasons. The repeated news release finally induces new momentum.[5]

Another possibility that can link news momentum to information environments is firms' strategic information disclosure. When insiders' disclosure preferences are not aligned with those of outside investors, various incentives can motivate managers to strategically disclose or withhold firm-specific information (e.g., Kothari, Shu, and Wysocki, 2009). For example, in fear of litigation, managers are willing to quickly reveal bad news (e.g., Skinner, 1994; Baginski, Hassell, and Kimbrough, 2002). Moreover, to lower the exercise price of their employee options, managers would accelerate the dissemination of bad news and withhold good news (e.g., Yermack, 1997; Aboody and Kasznik, 2000). In contrast to these incentives to disclose bad news early, managers also would like to withhold bad news under certain circumstances. For example, managers suffer a reduction in compensation and face high career concerns once bad news is released to the public (e.g., Nagar, Nanda, and Wysocki, 2003; Graham, Harvey, and Rajgopal, 2005; Kothari, Shu, and Wysocki, 2009). Taken together, these incentives may lead to the possibility that managers who disclose bad (good) news in the past may release more bad (good) news in the future.

[5] The promotion of financial news stories might cause biased information flow and the resulting news momentum. For example, Bushee and Miller (2007) and Solomon (2012) demonstrate that investor relations firms spin their clients' news, generating more media coverage of positive press releases than negative press releases. Similarly, Reuter and Zitzewitz (2006) and Gurun and Butler (2012) find that when local media report more news about local companies, they use fewer negative words compared to the same media reporting about nonlocal companies. They also document that a potential explanation for this positive slant is the firms' local media advertising. In particular, firms have incentive to promote news stories during major events (Ahern and Sosyura, 2014). If biased news releases persist for a while, it induces new momentum.

Because either stale news dissemination or strategic information disclosures is characterized by firms' information environments, we expect more stale news to be disseminated repeatedly or mass information to be withheld and released sequentially by managers in firms with poor information environments. This forms our *information environment hypothesis* for the understanding of news momentum:

H1: *The news momentum phenomenon is more pronounced in firms with poor information environments.*

Our second conjecture on what induces news momentum is related to firms' fundamentals. It has been well documented that earnings are persistent, and hence predictable (e.g., Ball and Brown, 1968; Beaver, Clarke and Wright, 1979; Fama and French, 2000, 2006; Markov and Tamayo, 2006; Li, 2010; Novy-Marx, 2015). In particular, Hou, van Dijk, and Zhang (2011) show that cross-sectional models are able to explain a large fraction of the variation in expected profitability across firms.

If firm-specific news stories fairly reflect firms' fundamentals, the persistent earning stream is likely to generate a positive correlation between sequential news releases. More specifically, positive news now is more likely to be followed by positive news in the future. Therefore, we propose *the fundamental hypothesis* to explain the news momentum phenomenon:

H2: *Firms with more positive news now will have better fundamentals in the future and this persistent difference in firm's fundamentals induces news momentum.*

Even though the economic mechanisms of the two hypotheses underlying news momentum are substantially different, we would like to emphasize that the two hypotheses are not mutually exclusive. For instance, it is likely that firms with the persistence of fundamentals attract media attention and have stale news broadcasted in the future.

3.3. Hypothesis tests

3.3.1. News momentum and information environments

To examine the information environment hypothesis, we follow the literature (e.g., Zhang, 2006) and adopt several firm-level characteristics to proxy for firms' information environments. These characteristics include firm size, analyst coverage, and institutional holdings. Firms with large market capitalization, high analyst coverage, and more institutional holdings tend to have more transparent information environments. If news momentum is driven by information environments, we would expect the news momentum phenomenon to be less significant for these transparent stocks. By looking at the magnitude of news momentum across groups sorted by firm size, analyst coverage, and institutional holding, it provides insights on the information environment story of news momentum.

Table 4 presents news momentum results by using the independent double sorting approach. Specifically, at the end of month *t*, we sort all stocks into five portfolios based on their news sentiment scores. We further independently sort all stocks into three portfolios (below the 30^{th} percentile and above the 70^{th} percentile) based on their previous year-end market capitalization (*Size*), analyst coverage (Analyst), or institutional holdings (*InstOwn*).

Panel A reports the results for independent double sorting according to firm size and news sentiment scores. It is evident that news momentum consistently presents in both the large *Size* tercile and the small *Size* tercile, confirming the robustness of news momentum. More importantly, if we look at the news sentiment difference between the GMB *News* portfolio of small firms and the GMB *News* portfolio of large firms, we find that there is no consistent difference in the news momentum pattern. For example, at the one-month horizon, large firms show slightly stronger news momentum. For the holding period from *t*+2 to *t*+6, however, small firms exhibit stronger news momentum, though the difference is only marginally significant. When the holding period is longer than half a year, the difference in news momentum between large and small firms becomes insignificant. These results do not provide support for the first hypothesis that firms' information environments cause news momentum.

[Insert Table 4 Here]

Panels B and C report the double sorting results for news momentum based on analyst coverage and institutional ownership. We again find insignificant difference in news momentum across stock portfolios with different levels of analyst coverage or institutional ownership. Intuitively, stocks with more analyst coverage or institutional holdings are more closely monitored by professional market participants. As such, firms with more analyst coverage or institutional holdings have a more transparent information environment. If new momentum is due to information environments, firms with more analyst coverage or institutional holdings should exhibit weaker news momentum. Our evidence is thus against the information environments hypothesis.

To summarize, our double sorting analysis leads to two conclusions: first, news momentum is a relatively robust phenomenon, which is not attenuated by firm characteristics such as firm size, analyst coverage, or institutional holdings; second, news momentum is unlikely to be driven by firms' information environments.

3.3.2. News momentum and firm fundamentals

If news momentum is driven by firms' fundamentals, we would expect current news to predict future firms' fundamentals. To test the *fundamental hypothesis*, we formally examine whether the news sentiment of the GMB portfolio contains information about firms' future fundamentals. Specifically, we use the return-on-assets ratio (*ROA*) and earnings surprise (*SUE*) as proxies for firms' fundamentals. Our earnings surprise measure is the equal-weighted firms' standardized unexpected earnings. In the spirit of Tetlock, Saar-Tsechansky, and Macskassy (2008), we compute each firm's *SUE* as

$$SUE_t = \frac{E_t - \mu_t}{\sigma_t}, \qquad (1)$$

where E_t is announced earnings, and μ_t and σ_t are the mean and standard deviation of forecasted earnings.

[Insert Table 5 Here]

Due to the standard point of data availability, we conduct our empirical tests on the relation between news sentiment and future *ROA* based on quarterly data. The interesting pattern is that current news stories indeed predict firms' future profitability. That is, firms with bad news stories have lower future *ROA*, while firms with good news releases generate more future profits. Combined together, we find that the GMB *News* portfolio implies a positive one-quarter-ahead *ROA* of 1.920 with a *t*-value of 6.55. The magnitude is also economically significant, which is equivalent to the 30.9% of the standard deviation of *ROA*. The results are robust to future *ROA*s in longer time horizons, implying the persistence of the predictive power of the GMB portfolio sentiment.

Panel B summarizes the results on the relationship between current news sentiment scores and future *SUE*. Consistent with the evidence on *ROA*, firms in the good (bad) *News* portfolio experience higher (lower) future *SUE* at various forecasting horizons. The differences in future *SUEs* between the good and bad *News* portfolios are all statistically significant at the 1% level. These results indicate that the GMB *News* portfolio persistently predicts the difference in firm's fundamentals, thus providing the supportive evidence for the *fundamental hypothesis*.[6]

To further pin down the fundamental explanation of news momentum, we divide news articles into two categories: hard news and soft news. The hard news group consists of four categories of news: "revenues", "earnings", "analyst-ratings", and "credit-ratings". All other news categories are included in the soft news group. Given hard news is relevant to firms' fundamentals by definition, we expect hard news to have a stronger effect in predicting firms' future fundamentals than soft news. To test this conjecture, we respectively use hard and soft news sentiments to form GMB *News* portfolios and compute the future *ROA* and *SUE* of GMB *News* portfolios.

[Insert Table 6 Here]

Table 6 presents the empirical results. In Panel A, we find that the GMB portfolio for hard news is associated with a positive ROA at the time of portfolio formation, and predicts a tendency of positive ROA in the future. This finding suggests that positive hard news implies firms' solid future fundamentals. In contrast to the result on hard news, firms with good (bad) soft news tend to have poorer (better) operational performance in the future. We obtain similar results when predicting future *SUE* in Panel B. Overall, these findings show that only hard news can significantly and positively forecast firms' future fundamentals.

[6] We also examine the fundamental prediction of news using the Fama-Macbeth (1973) approach by controlling for current *ROA* or *SUE* in the Internet Appendix Table IA2.

4. News momentum and return predictability

4.1. Baseline results

How news is incorporated into the stock price is central to the efficiency of the stock market. Given the cross-sectional pattern of news documented in Section 3, we would like to investigate whether investors are aware of news momentum. Specifically, we examine whether news sentiment scores predict cross-sectional future stock returns. If news momentum is correctly incorporated into the stock prices, stocks in the highest news score portfolio should have similar future returns as stocks in the lowest news score portfolio, and vice versa.

To test the pricing implication of news momentum, we form a news momentum trading strategy by buying stocks with high news sentiment scores and selling stocks with low news sentiment scores. Equivalent to the sorting strategy discussed in Section 3.1, at the end of month *t*, we sort our sample stocks into five portfolios based on news sentiment scores. We then hold the good *News* portfolio and sell the bad *News* portfolio. Last, we calculate the equally weighted future average returns for this GMB *News* portfolio. It is noteworthy that our news momentum trading strategy is totally different from the traditional momentum strategy (Jegadeesh and Titman, 1993): while our news momentum strategy sorts stocks based on news sentiment scores, the traditional strategy sorts stocks based on past performance.[7]

Table 7 reports the one-month-ahead portfolio returns obtained from the news momentum trading strategy. The first column indicates that there is significant news-driven return momentum: the strategy that buys the good *News* portfolio and sells the bad *News* portfolio generates a return of 0.696 percent per month (*t*-statistic=4.31), which is equivalent to a return of 8.352 percent per year. Look at all five portfolios, momentum profits rise monotonically.

[Insert Table 7 Here]

A major concern is whether the positive return of the news momentum trading strategy is obtained from its exposures to risk factors. To alleviate this concern, we respectively use the CAPM model, the Fama-French (1992) three factor model, the Fama-French-Cahart (Fama and French, 1993; Carhart, 1997) four factor model, and the Fama-French (2015) five factor model to

[7] Daniel and Moskowitz (2013) and Barroso and Santa-Clara (2015) find that the traditional return momentum strategy crashes during our sample period. Our news momentum strategy generates returns that cannot be explained by the traditional return momentum.

control for the risk exposures of GMB *News* portfolio profits. Specifically, we regress the excess returns of GMB *News* portfolios against the respective factors and calculate the regression intercepts that represent risk-adjusted returns, namely, alpha.

We know from Table 3 that the bad *News* portfolio has a negative news sentiment score at the end of month t. Interestingly, we find that this negative sentiment score leads to a negative return in one month after risk adjustment. In contrast, the raw return without adjusting for risk exposures is positive. More importantly, after risk adjustment, we find that the news momentum trading strategy even generates higher future stock returns. For example, the Fama-French three-factor model adjusted monthly return is 0.800, which is significant at the 1% level. After the adjustment of the Fama-French five factors, the alpha of the GMB *News* portfolio is 0.770, which is higher than the unadjusted raw return.

Overall, these findings suggest that stocks in the highest news score portfolio have higher future returns than stocks in the lowest news score portfolio, and the return profit of the news momentum trading strategy cannot be explained by its exposures to popularly used risk factors.[8]

4.2. Understanding return predictability

There are two explanations for the cross-sectional return predictability: mispricing and risk. The mispricing explanation is that investors are not aware of news momentum and underreact to news. In this case, news momentum would induce price continuation in the short run, but price continuation will not last for a long period, largely due to the reason that underreaction shocks are stationary and cannot last forever (e.g., Delong, Shleifer, Summers, and Waldmann, 1990). Moreover, the return predictability implied by news momentum should be stronger in firms with opaque information environments, where market participants are more likely to falsely make an inference.

On the other hand, rational asset pricing theory posits that stock return predictability can result from its exposures to time-varying aggregate risk, and to the extent that news sentiment consistently links to this time-varying aggregate risk premium, they will likely remain successful in return forecasts. If return momentum is caused by risk, we should observe that return momentum caused by news sentiment lasts for a long period and will not disappear quickly. In

[8] We examine the robustness of the return predictability of news by using different specifications in the Internet Appendix Table IA3.

addition, the return predictability induced by news momentum should not be affected by firms' information environments.

To shed light on the mispricing and risk-based explanations, we conduct two tests. First, we investigate the profitability of the news momentum strategy for longer holding periods. Specifically, we compute the equal-weighted Fama-French five-factor alphas for each portfolios and the GMB *News* portfolio with a holding period from *t*+2 and *t*+6, from *t*+7 to *t*+12, and from *t*+13 to *t*+24. The results are presented in Table 8.

[Insert Table 8 Here]

The strategy that buys the good *News* portfolio and sells the bad *News* portfolio generates a return of 0.243 percent per month with a *t*-statistic of 3.32 for the half year holding period. At longer horizons, news momentum profits disappear quickly. For the holding periods from t+7 to t+12 and from t+13 to t+24, the profit of this trading strategy is respectively -0.011 and -0.048 percent per month, which are economically and statistically insignificant. These results suggest that return predictability of news momentum is less likely to be explained by risk.

Second, we examine whether the trading profit of the news momentum strategy varies according to firms' information environments. We use firm size, analyst coverage, and institutional holdings as proxies for firms' information environments. At the end of each month *t*, we independently sort all stocks into five portfolios based on news sentiment scores and also three portfolios based on these firm characteristics. We then calculate the equal-weighted future stock returns for these portfolios.

Panel A of Table 9 reports the results for independent double sorting according to firm size and news sentiment scores. The return predictability implied by the news momentum is pronounced only in the small *Size* tercile but not in the large *Size* tercile. The raw return of the GMB *News* portfolio is 1.262 percent per month. After the adjustment for risk exposures to the Fama-French five factors, the next period return is 1.253 percent per month (*t*-statistic = 6.44). These results are consistent with the mispricing view of return predictability. Because small firms attract less attention and have less news releases, so information is likely to be more asymmetric and diffuses more slowly for these stocks. As such, we would expect to observe stronger trading profits for small firms. We find similar results for stocks with low analyst coverage in Panel B and stocks with less institutional ownership in Panel C.

[Insert Table 9 Here]

Taken together, these results suggest that the future return of the GMB *News* portfolio is only pronounced in short time horizons and in firms with opaque information environments, thus providing supporting evidence on the mispricing view of return predictability.

4.3. Future returns on news days and non-news days

Up to now, our analysis indicates that the news-driven momentum is attributable to mispricing, more specifically, underreaction. In the context of mispricing, there are two forms of underreaction that can account for observed return predictability. The first form of mispricing is that market participants do not realize the presence of news continuation and underestimate the persistence of news (fundamentals). The other form is that investors underreact to current news and induce return continuation such as the post-earnings announcement drift anomaly.[9] The two forms of mispricing represent two different economic channels through which news releases affect stock returns. While the first channel has never been examined, the second channel has been well-documented in the literature.

This section attempts to enrich our insights on the two economic channels. To accomplish this, we decompose future returns into returns on news days and non-news days. If news-driven return predictability is caused by the underestimation of news momentum, we should observe positive abnormal returns on news days, but not on non-news days. The intuition is that investors will realize underestimation only in subsequent news days. On the other hand, if news-driven return predictability is induced by underreaction to news at time *t*, we can only observe positive abnormal returns on non-news days. This is because subsequent stock prices will adjust to correctly reflect news sentiment at time *t* before next news releases.

Empirically, we decompose the return in a month next to the formation of the GMB portfolio into news-day returns and non-news-day returns. News-day returns (non-news-day returns) are the accumulative daily returns for all news days (non-news days) for a particular

9 On the theoretical side, the conservatism-bias model of Barberis, Shleifer, and Vishny (1998) and the heterogeneity model of Hong and Stein (1999) can account for the underreaction behavior of investors.

stock in a month. Table 10 presents the one-month-ahead portfolio returns of taking the news momentum trading strategy on news and non-news days, respectively. Panel A indicates that there is significant news-driven return momentum on news days: the strategy that buys the high-news-sentiment portfolio and sells the low-news-sentiment portfolio delivers a return of 0.324 percent per month with a t-value of 2.52 after the adjustment of the Fama-French five factors. These results are consistent with the channel that investors underestimate news momentum.

[Insert Table 10 Here]

Panel B reports the returns of the GMB trading strategy on non-news days. Similarly, we find the evidence on return continuation. Indeed, the news momentum trading strategy generates a return of 0.437 percent on non-news days after the adjustment of the Fama-French five factors. These results provide the supporting evidence on the economic channel that investors underreact to news sentiment at time t, which is associated with firms' fundamentals.

Overall, news-driven return predictability is caused by two economic channels. On the one hand, market participants do not realize the existence of news momentum and thus underreact to the persistence of news. On the other hand, investors underreact to news at time t and lead to return predictability.

4.4. Additional results

In light of the evidence regarding the information content of the hard and soft news categories, we analyze how soft and hard news releases affect return predictability. Specifically, we explore whether the effect of news sentiment on stock returns concentrates amongst some specific news categories. Toward this end, we regroup 36 news categories originally divided by RavenPack into a hard news categories and a soft news category.

[Insert Table 11 Here]

Table 11 confirms that hard news releases induce significant return momentum. Using the Fama-French five factor model to adjust returns, the alpha (the excess return) is 0.948 percent per month or 11.4 percent per annum. Turing to the soft-news-driven return momentum, we find that the alpha is 0.363 percent per month. The alpha difference between the hard-news GMB portfolio and the soft-news GMB portfolio is 0.584 percent per month with a t-value of 2.26. These results are consistent with the results of news momentum for both hard news and soft news reported in Table 6.

Given the endogeneity of news releases, our second method is to use the multivariate Fama and MacBeth (1973) regressions to check the effect of qualitative news on stock returns. Specifically, we run the following regressions:

$$R_{t+1} = a + b_1 News_t + \sum_{i=1}^{k} b_i Z_{i,t} \quad + \varepsilon_{t+1}, \qquad (2)$$

Where R_{t+1} is the stock return in month t, $News_t$ is news sentiment at time t, and $Z_{i,t}$ includes control variables observed at time t. We respectively use five stock returns as the dependent variable. They are respectively the raw return, the CAMP-adjusted return, the Fama-French 3-factor model adjusted return, the Carhart 4-factor model adjusted return, the Fama-French 5-factor model adjusted return. The control variables include the logarithm of market capitalization (*LogSize*), book-to-market ratio (B/M), market beta (*Beta*), idiosyncratic volatility (*IdioVol*), past two-month stock returns ($R_{t-3,t-2}$), past three-month stock returns($R_{t-6,t-4}$), past six-month stock returns ($R_{t-12,t-7}$), and Amihud's (2002) illiquidity measure (*Illiquidity*).

[Insert Table 12 Here]

Table 12 presents the multivariate Fama and MacBeth regression results. We confirm the influence of news sentiment on stock returns. The regressions consistently generate a positive slope coefficient (b_1), which are significant at the 1% level. These results are consistent with those from portfolio analysis. To illustrate the magnitude of news impact, column M1 indicates a slope coefficient of 0.625 with a t-statistic of 2.91. This implies that one unit increase in news sentiment predicts a rise of approximately 7.5% per annum in future returns. We also find that the coefficient of news sentiment is roughly stable across the five regressions. In summary, the Fama-Macbeth regressions provide further supporting evidence on the effect of news sentiment on stock returns.

5. Conclusions

The cross-sectional pattern of news is far less investigated. Using a comprehensive sample of firm-level news articles, we investigate the patterns of news releases. We find a strong cross-sectional news momentum phenomenon: firms with relatively higher current news sentiment scores are likely to have higher sentiment scores in the future; firms with relatively lower sentiment scores have lower sentiment scores in the future. New momentum is persistent and lasts up to more than two years. In light of news momentum, we explore what drives news momentum. We provide two hypotheses. The first hypothesis views news momentum as from opaque information environments. The second hypothesis argues that firms' fundamentals drive news momentum. A set of empirical tests provides the supporting evidence on the fundamental-driven news momentum.

We then investigate the asset pricing implications of news momentum. The empirical analysis shows that news releases induce significant return momentum. Two alternative explanations for return momentum is respectively a risk-based story and a mispricing interpretation. We design a set of tests and conduct the empirical exercise to show that mispricing or underreaction is the main driving force of news-based return predictability. To enhance our understanding on underreaction, we distinguish underreaction to current news releases from underreaction to news momentum. The empirical analysis shows that both channels play a role for understanding mispricing that induces return predictability.

References

Ahern Kenneth R. and Sosyura Denis, 2011, Who writes the news? Corporate press releases during merger negotiations, *Journal of Finance* 69, 241–291.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

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

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159–178.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barroso, Pedro, and Pedro Santa-Clara, 2015, Momentum has its moments. *Journal of Financial Economics* 116, 111–120.

Beaver, William H, Roger Clarke, and William F. Wright, 1979, The association between unsystematic security returns and the magnitude of earnings forecast errors, *Journal of Accounting Research* 17, 316–340.

Baginski, Stephen P., John M. Hassell, and Michael D. Kimbrough, 2002, The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in U.S. and Canadian markets, *The Accounting Review* 77, 25–50.

Bushee, Brian J., and Gregory S. Miller, 2007, Investor relations, firm visibility, and investor following, Working paper, The Wharton School, University of Pennsylvania.

Carhart, Mark M., 1997, On persistence in mutual fund performance, Journal of Finance 52, 57–82.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011. In search of attention. Journal of Finance 66, 1461–1499.

Dai, Lili, Jerry T. Parwada, and Bohui Zhang, 2015, The governance effect of the media's news dissemination role: Evidence from insider trading, *Journal of Accounting Research* 52, 331–366.

Dang, Tung Lam, Fariborz Moshirian, and Bohui Zhang, 2015, Commonality in news around the world. *Journal of Financial Economics* 116, 82–110.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyan, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

Daniel, Kent D., and Tobias J. Moskowitz, 2013, Momentum crashes. Working Paper, Columbia University.

Delong, Bradford J., Andrei Shleifer, Lawrence H. Summers, and Robert Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 489–519.

Engelberg, Joseph E., David R. McLean, and Jeffrey Pontiff, 2016, Anomalies and news, *Working paper*.

Fama, Eugene F., and Kennth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

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

Fama, Eugene F., and Kennth R. French, 2000, Forecasting profitability and earnings. *Journal of Business* 73, 161–175.

Fama, Eugene F., and Kennth R. French, 2006, The value premium and the CAPM. *Journal of Finance* 61, 2163–2185.

Fama, Eugene F., and Kennth R. French, 2015. A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607–636.

Fang, Lily H., and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *Journal of Finance* 64, 2023–2052.

Fang, Lily. H., Joel Peress, and Lu Zheng, 2014, Does media coverage of stocks affect mutual funds' trading and performance? *Review of Financial Studies* 27, 3441–3466.

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

Gurun, Umit G., and Alexander W. Butler, 2012, Don't believe the hype: Local media slant, local advertising, and firm value, *Journal of Finance* 67, 561–598.

Hillert, Alexander, Heiko Jacobs, and Sebastian Muller, 2014, Media makes momentum, *Review of Financial Studies* 27, 3467–3501.

Hirshlefer, David, and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Financial Economic* 36, 337–386.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hou, Kewei, MathijsA.vanDijk, and Yinglei Zhang, 2011, The implied cost of capital: A new approach, *Journal of Accounting and Economics* 53, 504–526.

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

Jiang, Hao, Sophia Zhengzi Li, and Hao Wang, 2015. News Momentum. Working Paper, Michigan State University.

Kelley, Eric K., and Paul C. Tetlock, 2013, How wise are crowds? Insights from retail orders and stock returns, *Journal of Finance* 68, 1229–1265.

Kolasinski, Adam C., Adam V. Reed, and Matthew C. Ringgenberg, 2013, A multiple lender approach to understanding supply and search in the equity lending market, *Journal of Finance* 68, 559–595.

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

Lesmond, David A., Michael J. Schill, and Chunsheng Zhou, 2004, The illusory nature of momentum profits, *Journal of Financial Economics* 71, 349–380.

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

Loughran, Tim, and Bill Mcdonald, 2014, Measuring readability in financial disclosures, *Journal of Finance* 69, 1643–1671.

Mclean, David R., and Jeffrey Pontiff, 2016, Does academic research destroy return predictability? *Journal of Finance* 71, 5–32.

Markov, Stanimir, and Ane Temayo, 2006, Predictability in financial analyst forecast errors: Learning or Irrationality? *Journal of Accounting Research* 44, 725–761.

Nagar, Venky, Dhananjay Nanda, and Peter Wysocki, 2003, Discretionary disclosure and stock-based incentives, *Journal of Accounting and Economics* 34, 283–309.

Novy-Marx, Robert, 2015. Fundamentally, momentum is fundamental momentum. Working Paper.

Reuter, Jonathan, and Eric Zitzewitz, 2006, Do ads influence editors? Advertising and bias in the financial media, *Quarterly Journal of Economics* 121, 197–227.

Sagi, Jacob S., and Mark S. Seasholes, 2007, Firm-specific attributes and the cross-section of momentum, *Journal of Financial Economics* 84, 389–434.

Shroff, Nemit, Rodrigo S. Verdi, and Gwen Yu, 2014, Information environment and the investment decisions of multinational corporations, *Accounting Review* 89, 759–790.

Skinner, Douglas J., 1994, Why firms voluntarily disclose bad news? *Journal of Accounting Research* 32, 38–60.

Solomon, David H., 2012. Selective publicity and stock prices. *Journal of Finance* 67, 599–637.

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., 2010. Does public financial news resolve asymmetric information, *Review of Financial Studies* 23, 3520–3557.

Tetlock, Paul C., 2011, All the news that’s tit to reprint: Do investors react to stale information? *Review of Financial Studies* 24, 1481–1512.

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.

Veldkamp, Laura L., 2006, Information markets and comovement of asset prices, *Review of Economic Studies* 73, 823–845.

Yermack, David, 1997, Good timing: CEO stock option awards and company news announcements. *Journal of Finance* 52, 449–476.

Zhang, Frank, 2006, Information uncertainty and stock returns. *Journal of Finance* 61, 105–136.

Appendix A: Definitions of the Variables

Variable	Acronym	Definition	Source
News	$News_t$	Average ESS score of all news for a particular firm over a month (quarter/week/day) t.	RavenPack
Hard news	$HardNews_t$	Average ESS score of hard news for a particular firm over a month (quarter) t.	RavenPack
Soft news	$SoftNews_t$	Average ESS score of soft news for a particular firm over a month (quarter) t.	RavenPack
Next month returns	R_{t+1}	Stock return in percentage in month t+1.	CRSP
Market capitalization	$Size_t$	Market capitalization at the end of previous year.	CRSP
Analyst coverage	$Analyst_t$	Number of analysts following in month t.	IBES
Institutional ownership	$InstOwn_t$	Number of shares held by institutional investors divided by total shares outstanding in the previous quarter.	Thomson Reuters
ROA	ROA_t	The ratio of net income in quarter t over total assets in quarter t-1, which is scaled by 100 in the analysis.	Compustat
Earnings surprise	SUE_t	Earning surprise (SUE score) in quarter t.	IBES
Book-to-market ratio	B/M_t	The ratio of book value of equity to market value of equity in the previous year, which is winsorized at 1% and 99% cutoffs.	Compustat, CRSP
Market beta	$Beta_t$	Regression of $r_i=alpha+beta^{*}r_m+e$ from month t-59 to t.	CRSP
AHXZ's idiosyncratic volatility	$IdioVol_t$	Standard deviation of residuals from regression of $r_i=alpha+b_1^{*}(r_m-r_f)+b_2^{*}SMB+b_3^{*}HML+e$ over previous year by using daily returns.	CRSP, Fama & French
Past two-month stock returns	$\text{Return}_{t-3,t-2}$	Compounded return in percentage from month t-3 to t-2.	CRSP
Past three-month stock returns	$\text{Return}_{t-6,t-4}$	Compounded return in percentage from month t-6 to t-4.	CRSP
Past six-month stock returns	$\text{Return}_{t-12,t-7}$	Compounded return in percentage from month t-12 to t-7.	CRSP
Amihud's (2002) illiquidity	$Illiquidity_t$	Illiquidity is the daily ratio of absolute stock return to its dollar volume, averaged over previous year, which is scaled by 10,000 in the analysis.	CRSP

Appendix B: List of News by Categories

News Categories	News Groups	Frequency
Hard news	Revenues	5.96%
	Earnings	19.55%
	Analyst-ratings	3.57%
	Credit-ratings	1.26%
	Subtotal	**30.35%**
Soft news	Acquisitions-mergers	3.21%
	Assets	1.42%
	Balance-of-payments	0.00%
	Bankruptcy	0.04%
	Civil-unrest	0.00%
	Corporate-responsibility	0.04%
	Credit	0.82%
	Crime	0.00%
	Dividends	2.56%
	Equity-actions	2.88%
	Exploration	0.02%
	Government	0.01%
	Indexes	0.03%
	Industrial-accidents	0.01%
	Insider-trading	14.07%
	Investor-relations	5.29%
	Labor-issues	5.63%
	Legal	0.99%
	Marketing	3.44%
	Order-imbalances	6.99%
	Partnerships	1.41%
	Pollution	0.00%
	Price-targets	0.23%
	Products-services	8.47%
	Public-opinion	0.00%
	Regulatory	0.29%
	Security	0.01%
	Stock-prices	4.36%
	Taxes	0.00%
	Technical-analysis	7.42%
	Transportation	0.00%
	War-conflict	0.01%
	Subtotal	**69.65%**

Table 1: Summary Statistics

This table presents the summary statistics of main variables used in this study. The variables include news (*News*), hard news (*HardNews*), soft news (*SoftNews*), next month returns (R_{t+1}), logarithm of market capitalization (*LogSize*), analyst coverage (*Analyst*), institutional ownership (*InstOwn*), book-to-market ratio (*B/M*), beta (*Beta*), idiosyncratic volatility (*IdioVol),* past two-month stock returns ($Return_{t-3,\ t-2}$), past three-month stock returns ($Return_{t-6,\ t-4}$), past six-month stock returns ($Return_{t-12,\ t-7}$), Amihud's (2002) illiquidity (*Illiquidity*), earnings surprise (SUE_t) and ROA (ROA_t). All the variables are defined in Appendix A. The table reports the number of observations (*NObs*), mean, median, standard deviation (*STD*), quartile (*75% and 25%*), and the bottom/top 5% (*5% and 95%*) distribution of the variables. The sample period is from January 2000 to October 2014 and observations with zero news scores are not included.

Variable	NObs	Mean	STD	5%	25%	Median	75%	95%
$News_t$	473,941	0.083	0.157	-0.172	-0.012	0.083	0.181	0.332
$HardNews_t$	289,870	0.109	0.260	-0.323	-0.066	0.117	0.286	0.527
$SoftNews_t$	399,291	0.080	0.151	-0.157	-0.015	0.073	0.176	0.323
R_{t+1}	473,941	1.004	13.750	-18.482	-5.908	0.397	6.897	22.044
$Size_t$	473,941	6.361	1.946	3.242	5.007	6.309	7.625	9.723
$Analyst_t$	473,941	6.932	6.916	0.000	1.517	4.865	10.287	21.191
$InstOwn_t$	473,941	0.568	0.289	0.061	0.339	0.615	0.800	0.965
B/M_t	473,941	0.696	0.585	0.116	0.316	0.555	0.888	1.756
$Beta_t$	423,821	1.178	0.823	0.175	0.596	1.013	1.587	2.765
$IdioVol_t$	473,881	0.029	0.017	0.012	0.018	0.025	0.036	0.059
$Return_{t-3,\ t-2}$	472,750	3.062	21.360	-24.667	-7.915	1.273	11.177	35.583
$Return_{t-6,\ t-4}$	470,391	4.480	27.048	-29.246	-9.475	1.793	14.272	45.338
$Return_{t-12,\ t-7}$	461,523	9.845	43.107	-37.813	-11.938	3.997	22.585	74.125
$Illiquidity_t$	473,876	0.040	0.303	0.000	0.000	0.000	0.002	0.131
SUE_t	330,298	0.855	7.365	-4.232	-0.491	0.751	2.367	6.924
ROA_t	458,926	-0.203	6.714	-9.357	-0.236	0.669	1.926	4.689

Table 2: Number of News Articles per Month over Time

This table presents the number of news articles per month across different size groups over five time periods including 2000-2002, 2003-2005, 2006-2008, 2009-2011, and 2012-2014. Each month, firms are classified into 5 groups based on previous year end market capitalization (*Size*). Panel A reports the average number of all news articles per month. Panel B reports the average number of positive news articles per month. Panel C reports the average number of negative news articles per month. The sample period is from January 2000 to October 2014.

Panel A: The Number of All News Articles for Each Month

Portfolios	2000-2002	2003-2005	2006-2008	2009-2011	2012-2014
Small *Size*	2.28	3.95	5.31	5.92	7.33
2	2.84	5.10	6.63	7.82	11.74
3	3.36	6.13	8.12	9.94	15.69
4	4.05	7.35	10.19	12.80	20.05
Large *Size*	8.71	14.26	20.98	25.88	35.63

Panel B: The Number of Positive News Articles for Each Month

Portfolios	2000-2002	2003-2005	2006-2008	2009-2011	2012-2014
Small *Size*	1.07	1.69	2.36	2.35	3.03
2	1.30	2.02	2.73	2.95	4.67
3	1.53	2.28	3.22	3.68	6.05
4	1.91	2.89	4.20	5.04	8.10
Large *Size*	4.64	6.62	10.04	12.85	17.27

Panel C: The Number of Negative News Articles for Each Month

Portfolios	2000-2002	2003-2005	2006-2008	2009-2011	2012-2014
Small *Size*	0.57	0.86	1.04	1.10	1.51
2	0.65	1.19	1.44	1.68	3.06
3	0.76	1.59	1.94	2.41	4.63
4	0.96	1.89	2.51	3.26	6.22
Large *Size*	2.18	3.19	5.28	6.88	11.09

Table 3: Momentum of News

This table presents the momentum effects of news. At the end of month t, we sort all stocks with non-zero news scores into five portfolios based on their news scores ($News_t$). Stocks in *Bad News* portfolio have the lowest news scores while stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average news scores of each portfolio over different time periods after the portfolio formation. $News_{t+1}$ shows the 1-month average news scores of each portfolio in month t+1; $News_{t+2,\ t+6}$ shows the average news scores over 5 months from t+2 to t+6; $News_{t+7,\ t+12}$ shows the average news scores over 6 months from t+7 to t+12; and $News_{t+13,\ t+24}$ shows average news scores over 12 months from t+13 to t+24. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Portfolios	$News_t$	$News_{t+1}$	$News_{t+2,\ t+6}$	$News_{t+7,\ t+12}$	$News_{t+13,\ t+24}$
Bad *News*	-0.134	0.042	0.038	0.044	0.049
2	0.009	0.056	0.054	0.054	0.055
3	0.083	0.064	0.065	0.063	0.062
4	0.159	0.072	0.075	0.069	0.066
Good *News*	0.296	0.078	0.083	0.074	0.069
Good-Bad	**0.431*****$^{}$	**0.037*****	**0.044*****	**0.031*****	**0.020*****
	(26.41)	**(25.68)**	**(49.80)**	**(33.93)**	**(24.72)**

Table 4: Momentum of News by Different Information Environments

This table presents the momentum effects of news by different information environments. At the end of month t, we sort all stocks with non-zero news scores into five portfolios based on their news scores ($News_t$). We further independently sort all stocks into three portfolios based on their previous year end market capitalization (*Size*), analyst coverage (*Analyst*), and institutional ownership (*InstOwn*), respectively. Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average news scores of the "*Good-Bad*" portfolios for *Small/Large Size* subsamples, *Low/High Analyst* subsamples, *Low/High InstOwn* subsamples, as well as *"Small-Large" Size*, *"Low-High" Analyst* and *"Low-High" InstOwn* hedge portfolios over different time periods after the portfolio formation. $News_{t+1}$ shows the 1-month average news scores of each portfolio in month t+1; $News_{t+2,\ t+6}$ shows the average news scores over 5 months from t+2 to t+6; $News_{t+7,\ t+12}$ shows the average news scores over 6 months from t+7 to t+12; and $News_{t+13,\ t+24}$ shows average news scores over 12 months from t+13 to t+24. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Panel A: Momentum of News across Size Subsamples

Portfolios	$News_t$	$News_{t+1}$	$News_{t+2,\ t+6}$	$News_{t+7,\ t+12}$	$News_{t+13,\ t+24}$
Small *Size*	0.460***	0.034***	0.046***	0.031***	0.020***
	(27.70)	(24.25)	(42.35)	(29.84)	(22.58)
Large *Size*	0.393***	0.042***	0.043***	0.031***	0.022***
	(24.41)	(24.07)	(33.74)	(23.30)	(18.12)
Small-Large	0.068***	-0.008***	0.003*	0.000	-0.002
	(20.06)	(-4.32)	(1.79)	(-0.33)	(-1.60)

Panel B: Momentum of News across Analyst Coverage Subsamples

Portfolios	$News_t$	$News_{t+1}$	$News_{t+2,\ t+6}$	$News_{t+7,\ t+12}$	$News_{t+13,\ t+24}$
Low *Analyst*	0.459***	0.034***	0.046***	0.030***	0.019***
	(27.41)	(20.55)	(46.93)	(27.24)	(20.42)
High *Analyst*	0.394***	0.042***	0.045***	0.033***	0.023***
	(24.58)	(22.83)	(31.54)	(23.38)	(14.96)
Low-High	0.064***	-0.008***	0.000	-0.004**	-0.003**
	(19.54)	(-4.04)	(0.25)	(-2.40)	(-2.26)

Panel C: Momentum of News across Institutional Holdings Subsamples

Portfolios	$News_t$	$News_{t+1}$	$News_{t+2,\ t+6}$	$News_{t+7,\ t+12}$	$News_{t+13,\ t+24}$
Low *InstOwn*	0.458***	0.033***	0.046***	0.030***	0.020***
	(27.88)	(17.46)	(47.63)	(28.39)	(21.01)
High *InstOwn*	0.407***	0.042***	0.044***	0.033***	0.022***
	(25.22)	(27.22)	(35.14)	(32.46)	(22.43)
Low-High	0.051***	-0.009***	0.002	-0.003**	-0.003**
	(16.17)	(-4.04)	(1.39)	(-1.98)	(-2.13)

Table 5: News and Firm Fundamentals

This table examines the relation between the news and firm fundamentals. At the end of quarter t, we sort all stocks with non-zero news scores into five portfolios based on their news scores ($News_t$). Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average ROA (SUE) of each portfolio over different time periods after the portfolio formation. ROA_{t+1} (SUE_{t+1}) shows the 1-quarter average ROA (SUE) of each portfolio in quarter t+1; ROA_{t+2} (SUE_{t+2}) shows the 1-quarter average ROA (SUE) in quarter t+2; $ROA_{t+3,t+4}$ ($SUE_{t+3,t+4}$) shows the average ROA (SUE) over 2 quarters from t+3 to t+4; and $ROA_{t+5,t+8}$ ($SUE_{t+5,t+8}$) shows the average ROA (SUE) over 4 quarters from t+5 to t+8. Panel A reports the average *future ROA* for portfolios formed based on news ($News_t$). Panel B reports the average *future SUE* for portfolios formed based on news ($News_t$). The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Panel A: Future ROA

Portfolios	ROA_t	ROA_{t+1}	ROA_{t+2}	$ROA_{t+3,\ t+4}$	$ROA_{t+5,\ t+8}$
Bad *News*	-1.825	-1.756	-1.776	-1.748	-1.521
2	-0.592	-0.597	-0.630	-0.639	-0.477
3	0.034	-0.067	-0.083	-0.128	-0.098
4	0.319	0.210	0.077	0.035	0.055
Good *News*	0.250	0.164	0.068	0.010	-0.032
Good-Bad	**2.074*****	**1.920*****	**1.843*****	**1.758*****	**1.488*****
	(6.60)	**(6.55)**	**(6.77)**	**(6.53)**	**(6.55)**

Panel B: Future SUE

Portfolios	SUE_t	SUE_{t+1}	SUE_{t+2}	$SUE_{t+3,\ t+4}$	$SUE_{t+5,\ t+8}$
Bad *News*	-1.014	0.035	0.202	0.336	0.650
2	0.373	0.658	0.657	0.763	0.839
3	1.110	0.932	1.023	0.989	0.985
4	1.446	1.034	0.902	0.920	1.021
Good *News*	1.828	1.157	0.967	1.014	0.916
Good-Bad	**2.841*****	**1.122*****	**0.765*****	**0.678*****	**0.266*****
	(12.53)	**(7.26)**	**(8.54)**	**(5.91)**	**(4.33)**

Table 6: Hard News vs. Soft News and Firm Fundamentals

This table compares the firm fundamental difference between hard news and soft news. At the end of quarter t, we sort all stocks with non-zero news scores into five portfolios based on hard news (*Hard News*) and soft news (*Soft News*), respectively. Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted ROA and earnings surprise of the "*Good-Bad*" portfolios for *Hard News* portfolio, *Soft News* portfolio as well as "*Hard-Soft*" hedge portfolio which is long in Hard News and short in Soft News. Panel A (Panel B) reports the future ROA (earnings surprise) for hard and soft news. ROA_{t+1} (SUE_{t+1}) shows the 1-quarter average ROA (SUE) of each portfolio in quarter t+1; ROA_{t+2} (SUE_{t+2})shows the 1-quarter average ROA (SUE)) in quarter t+2; $ROA_{t+3,t+4}$ ($SUE_{t+3,t+4}$)shows the average ROA (SUE) over 2 quarters from t+3 to t+4; and $ROA_{t+5,t+8}$ ($SUE_{t+5,t+8}$)shows the average ROA (SUE) over 4 quarters from t+5 to t+8. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Panel A: Future ROA for Hard and Soft news

News Categories	ROA_t	ROA_{t+1}	ROA_{t+2}	$ROA_{t+3,\, t+4}$	$ROA_{t+5,\, t+8}$
Hard News	4.250***	3.993***	3.850***	3.687***	3.206***
	(20.79)	(18.03)	(20.41)	(19.72)	(20.92)
Soft News	-1.369***	-1.523***	-1.544***	-1.490***	-1.387***
	(-8.46)	(-13.08)	(-12.67)	(-11.48)	(-10.92)
Hard - Soft	5.620***	5.516***	5.394***	5.177***	4.594***
	(39.22)	(31.57)	(40.42)	(37.13)	(31.77)

Panel B: Future SUE for Hard and Soft news

News Categories	SUE_t	SUE_{t+1}	SUE_{t+2}	$SUE_{t+3,\, t+4}$	$SUE_{t+5,\, t+8}$
Hard News	4.166***	1.626***	1.279***	0.988***	0.501***
	(12.74)	(7.62)	(10.09)	(6.36)	(5.61)
Soft News	-0.229*	-0.012	-0.143	-0.057	-0.104
	(-1.93)	(-0.12)	(-0.84)	(-0.76)	(-1.44)
Hard-Soft	4.395***	1.638***	1.422***	1.044***	0.605***
	(11.34)	(6.16)	(5.78)	(5.14)	(4.53)

Table 7: Return Predictability of News

This table presents the return predictability of news by examining the average next month returns of portfolios constructed based on monthly news scores. At the end of month t, we sort all stocks with non-zero news scores into five portfolios based on news scores ($News_t$). Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted one-month-ahead average return (R_{t+1}), CAPM alpha ($R_{CAPM,\, t+1}$), Fama and French three factor alpha ($R_{FF3,\, t+1}$), four factor alpha ($R_{FF4,\, t+1}$) and five factor alpha ($R_{FF5,\, t+1}$) for each portfolio. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Portfolios	R_{t+1}	$R_{CAPM,\, t+1}$	$R_{FF3,\, t+1}$	$R_{FF4,\, t+1}$	$R_{FF5,\, t+1}$
Bad *News*	0.642	-0.001	-0.233	-0.153	-0.189
2	0.909	0.294	0.045	0.103	-0.001
3	0.981	0.399	0.176	0.215	0.143
4	1.142	0.561	0.329	0.355	0.319
Good *News*	1.337	0.776	0.568	0.602	0.581
Good-Bad	**0.696*** **	**0.777*** **	**0.800*** **	**0.755*** **	**0.770*** **
	(4.31)	**(5.22)**	**(5.26)**	**(5.40)**	**(4.88)**

Table 8: Return Predictability of News for Different Time Horizon

This table presents the return predictability of news for different time horizon by examining the average monthly Fama-French five factor alphas of portfolios constructed based on monthly news scores for different holdings periods. At the end of month t, we sort all stocks with non-zero news scores into five portfolios based on their news scores ($News_t$). Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average monthly Fama and French five factor alpha for different holdings periods after the portfolio formation. $R_{FF5,\, t+1}$ shows the 1-month average FF alpha in month t+1; $R_{FF5,\, t+2,\, t+6}$ shows the average FF alpha over 5 months from t+2 to t+6; $R_{FF5,\, t+7,\, t+12}$ shows the average FF alpha over 6 months from t+7 to t+12; and $R_{FF5,\, t+13,\, t+24}$) shows the average FF alpha over 12 months from t+13 to t+24. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Portfolios	$R_{FF5,\, t+1}$	$R_{FF5,\, t+2,\, t+6}$	$R_{FF5,\, t+7,\, t+12}$	$R_{FF5,\, t+13,\, t+24}$
Bad *News*	-0.189	0.219	0.406	0.424
2	-0.001	0.176	0.289	0.270
3	0.143	0.142	0.270	0.233
4	0.319	0.270	0.298	0.241
Good *News*	0.581	0.462	0.394	0.375
Good-Bad	**0.770*****	**0.243*****	**-0.011**	**-0.048**
	(4.88)	**(3.32)**	**(-0.20)**	**(-1.37)**

Table 9: Return Predictability of News by Different Information Environments

This table presents the return predictability of news in different information environments. At the end of month t, we sort all stocks with non-zero news scores into five portfolios based on their news scores ($News_t$). We further independently sort all stocks into three portfolios based on their previous year end market capitalization (*Size*), analyst coverage (*Analyst*), and institutional ownership (*InstOwn*), respectively. Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average returns of the "*Good-Bad*" portfolios for *Small/Large Size* subsamples, *Low/High Analyst* subsamples, *Low/High InstOwn* subsamples, as well as *"Small-Large" Size*, *"Low-High" Analyst* and *"Low-High" InstOwn* hedge portfolios over different time periods after the portfolio formation. Return measures include equally weighted one-month-ahead average return (R_{t+1}), CAPM alpha ($R_{CAPM,\ t+1}$), Fama and French three factor alpha ($R_{FF3,\ t+1}$), four factor alpha ($R_{FF4,\ t+1}$) and five factor alpha ($R_{FF5,\ t+1}$) for each portfolio. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Panel A: Return Predictability of News for Size Subsamples

Portfolios	R_{t+1}	$R_{CAPM,\ t+1}$	$R_{FF3,\ t+1}$	$R_{FF4,\ t+1}$	$R_{FF5,\ t+1}$
Small *Size*	1.168***	1.261***	1.274***	1.232***	1.253***
	(5.74)	(6.87)	(6.78)	(6.85)	(6.44)
Large *Size*	0.065	0.120	0.190	0.129	0.228
	(0.36)	(0.65)	(1.04)	(0.79)	(1.21)
Small-Large	1.103***	1.142***	1.084***	1.103***	1.025***
	(4.64)	(5.26)	(4.94)	(5.04)	(4.55)

Panel B: Return Predictability of News for Analyst Coverage Subsamples

Portfolios	R_{t+1}	$R_{CAPM,\ t+1}$	$R_{FF3,\ t+1}$	$R_{FF4,\ t+1}$	$R_{FF5,\ t+1}$
Low *Analyst*	1.172***	1.253***	1.272***	1.254***	1.244***
	(6.18)	(7.83)	(7.81)	(7.75)	(7.37)
High *Analyst*	-0.007	0.058	0.087	0.015	0.066
	(-0.03)	(0.29)	(0.43)	(0.08)	(0.32)
Low-High	1.178***	1.194***	1.185***	1.239***	1.177***
	(4.80)	(5.09)	(4.95)	(5.42)	(4.77)

Panel C: Return Predictability of News for Institutional Holdings Subsamples

Portfolios	R_{t+1}	$R_{CAPM,\ t+1}$	$R_{FF3,\ t+1}$	$R_{FF4,\ t+1}$	$R_{FF5,\ t+1}$
Low *InstOwn*	1.199***	1.293***	1.271***	1.229***	1.147***
	(5.68)	(6.64)	(6.46)	(6.50)	(5.71)
High *InstOwn*	0.065	0.100	0.175	0.110	0.253
	(0.38)	(0.56)	(1.00)	(0.72)	(1.40)
Low-High	1.135***	1.194***	1.096***	1.119***	0.894***
	(4.42)	(5.18)	(4.92)	(5.06)	(4.01)

Table 10: News-day Returns and Non-news-day Returns

This table presents the return predictability of news by separating the next month returns into news-day returns and non-news-day return. If news is reported in day t, then day t-1, t, and t+1 are treated as news days. Days without news are defined as non-news days. News-day returns (Non-news-day returns) are the accumulative daily returns for all news days (non-news days) for a particular stock in a month. At the end of month t, we sort all stocks with non-zero news scores into five portfolios based on news scores ($News_t$). Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted one-month-ahead average return (R_{t+1}), CAPM alpha ($R_{CAPM,\,t+1}$), Fama and French three factor alpha ($R_{FF3,\,t+1}$), four factor alpha ($R_{FF4,\,t+1}$) and five factor alpha ($R_{FF5,\,t+1}$) for each portfolio. Panel A reports the news-day returns and Panel B reports the non-news day returns. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Panel A: News-day Returns

Portfolios	R_{t+1}	$R_{CAPM,\,t+1}$	$R_{FF3,\,t+1}$	$R_{FF4,\,t+1}$	$R_{FF5,\,t+1}$
Bad *News*	0.673	0.323	0.177	0.200	0.130
2	0.743	0.359	0.208	0.233	0.125
3	0.741	0.372	0.243	0.257	0.173
4	0.837	0.484	0.361	0.370	0.309
Good *News*	0.883	0.577	0.482	0.488	0.454
Good-Bad	**0.211****	**0.254*****	**0.305*****	**0.288*****	**0.324*****
	(2.52)	**(3.46)**	**(4.20)**	**(4.16)**	**(4.29)**

Panel B: Non-news-day Returns

Portfolios	R_{t+1}	$R_{CAPM,\,t+1}$	$R_{FF3,\,t+1}$	$R_{FF4,\,t+1}$	$R_{FF5,\,t+1}$
Bad *News*	0.015	-0.429	-0.499	-0.443	-0.416
2	0.198	-0.186	-0.272	-0.239	-0.242
3	0.292	-0.072	-0.154	-0.129	-0.124
4	0.339	-0.040	-0.134	-0.117	-0.096
Good *News*	0.492	0.087	-0.015	0.013	0.021
Good-Bad	**0.478*****	**0.516*****	**0.484*****	**0.457*****	**0.437*****
	(3.65)	**(4.33)**	**(3.99)**	**(3.94)**	**(3.50)**

Table 11: Return Predictability of Hard News vs. Soft News

This table compares the difference of return predictability between hard news and soft news. At the end of month (or quarter) t, we sort all stocks with non-zero news scores into five portfolios based on hard news (*Hard News*) and soft news (*Soft News*), respectively. Stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average returns of the "*Good-Bad*" portfolios for *Hard News* portfolio, *Soft News* portfolio as well as "*Hard-Soft*" hedge portfolio which is long in Hard News and short in Soft News. Return measures include equally weighted one-month-ahead average return (R_{t+1}), CAPM alpha ($R_{CAPM,\, t+1}$), Fama and French three factor alpha ($R_{FF3,\, t+1}$), four factor alpha ($R_{FF4,\, t+1}$) and five factor alpha ($R_{FF5,\, t+1}$) for each portfolio. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

News Categories	R_{t+1}	$R_{CAPM,\, t+1}$	$R_{FF3,\, t+1}$	$R_{FF4,\, t+1}$	$R_{FF5,\, t+1}$
Hard News	0.750***	0.861***	0.950***	0.836***	0.948***
	(2.89)	(3.78)	(4.12)	(4.86)	(3.96)
Soft News	0.205*	0.219*	0.273**	0.277**	0.364***
	(1.77)	(1.78)	(2.42)	(2.45)	(3.19)
Hard-Soft	0.544*	0.642**	0.677***	0.559***	0.584**
	(1.80)	(2.55)	(2.70)	(2.88)	(2.26)

Table 12: Return Predictability of News—Fama-MacBeth Regressions

This table presents Fama-MacBeth Regressions of next month returns or alphas on news scores (*News*$_t$) and control variables. The dependent variables are stock returns in M1 (R_{t+1}), CAPM alphas in M2 ($R_{CAPM,\ t+1}$), Fama and French three factor alphas in M3 ($R_{FF3,\ t+1}$), four factor alphas in M4 ($R_{FF4,\ t+1}$) and five factors alphas in M5 ($R_{FF5,\ t+1}$). The control variables include ROA (*ROA*$_t$), earnings surprise (*SUE*$_t$), logarithm of market capitalization (*LogSize*), book-to-market ratio (*B/M*), beta (*Beta*), idiosyncratic volatility (*IdioVol)*, past two-month stock returns (*Return*$_{t-3,\ t-2}$), past three-month stock returns (*Return*$_{t-6,\ t-4}$), past six-month stock returns (*Return*$_{t-12,\ t-7}$), and Amihud's (2002) illiquidity (*Illiquidity*). All the variables are defined in Appendix A. The table also reports the number of observations (*NObs*), number of firms (*Firms*), and adjusted R square (*Adj-R*^{2}). The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

	R_{t+1}	$R_{CAPM,\ t+1}$	$R_{FF3,\ t+1}$	$R_{FF4,\ t+1}$	$R_{FF5,\ t+1}$
Variable	M1	M2	M3	M4	M5
News$_t$	0.625***	0.558***	0.644***	0.675***	0.685***
	(2.91)	(2.68)	(3.38)	(3.70)	(3.39)
ROA$_t$	0.053***	0.053***	0.055***	0.054***	0.044***
	(3.85)	(3.93)	(4.33)	(4.32)	(3.65)
SUE$_t$	-0.002	-0.003	-0.003	-0.004	-0.001
	(-0.38)	(-0.47)	(-0.64)	(-0.70)	(-0.12)
Return$_{t-3,\ t-2}$	-0.001	-0.000	-0.001	-0.000	-0.001
	(-0.24)	(-0.09)	(-0.12)	(-0.02)	(-0.33)
Return$_{t-6,\ t-4}$	-0.003	-0.003	-0.002	-0.003	-0.004
	(-0.61)	(-0.65)	(-0.63)	(-0.75)	(-1.16)
Return$_{t-12,\ t-7}$	-0.003	-0.002	-0.002	-0.002	-0.002
	(-1.06)	(-1.04)	(-1.03)	(-0.73)	(-1.07)
Size$_t$	-0.148***	-0.129***	-0.065**	-0.061**	-0.044
	(-3.06)	(-2.65)	(-2.14)	(-2.00)	(-1.60)
B/M$_t$	0.104	0.189	0.068	0.071	-0.047
	(0.68)	(1.23)	(0.62)	(0.68)	(-0.41)
Beta$_t$	-0.021	-0.289*	-0.202	-0.144	-0.172
	(-0.11)	(-1.80)	(-1.36)	(-1.05)	(-1.10)
IdioVol$_t$	-7.810	-6.228	-3.808	-5.586	7.871
	(-0.76)	(-0.74)	(-0.50)	(-0.75)	(1.12)
Illiquidity$_t$	-5.174	-5.071	-0.424	0.278	-0.752
	(-0.69)	(-0.70)	(-0.06)	(0.04)	(-0.11)
Intercept	1.932***	1.636***	0.792**	0.780**	0.354
	(3.67)	(3.01)	(2.00)	(2.05)	(0.99)
NObs	290,228	290,228	290,228	290,228	290,228
Firms	4,381	4,381	4,381	4,381	4,381
Adj-R^2	0.080	0.059	0.041	0.038	0.042

Table IA1: Momentum of News—Robustness Tests

The table examines the robustness of momentum effects of news by using different specifications. "*Aggregate News*" means that each month all stocks are grouped into five portfolios based on their aggregate news scores (*News*$_t$), which is the sum of ESS score of all news for a particular firm over a month t. "*News-day returns*" means that each month all stocks are grouped into five portfolios based on their News-day returns which is the average 3-day returns (-1, 1) around each news for all news for a particular firm over a month t. *"Daily"* and *"Weekly"* means that all stocks are grouped into five portfolios based on their news scores (*News*$_t$) at the end of day t and week t, respectively. *"Neutral News Included"* means that all stocks including those with zero news scores (neutral news or no news) are grouped into five portfolios based on their news scores (*News*$_t$) at the end of month t. *"Decile Portfolios"* means that all stocks are grouped into ten portfolios based on news scores (*News*$_t$) at the end of month t. For these four specifications, stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. For the specification *"Negative vs. Positive"*, stocks in *Bad News* portfolio have the negative news scores and stocks in *Good News* portfolio have the positive news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted average news scores of *"Good-Bad"* hedge portfolio over different time periods after the portfolio formation. *News*$_{t+1}$ shows the average news scores in month (or week, or day) t+1; *News*$_{t+2,\, t+6}$ shows the average news scores over 5 months from t+2 to t+6 (or 4 days from t+2 to t+5, or 3 weeks from t+2 to t+4); *News*$_{t+7,\, t+12}$ shows the average news scores over 6 months from t+7 to t+12 (or 5 days from t+6 to t+10, or 8 weeks from t+5 to t+12); and *News*$_{t+13,\, t+24}$ shows average news scores over 12 months from t+13 to t+24 (or 10 days from t+11 to t+20, or 12 weeks from t+13 to t+24). The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Specifications	*News*$_t$	*News*$_{t+1}$	*News*$_{t+2,\, t+6}$	*News*$_{t+7,\, t+12}$	*News*$_{t+13,\, t+24}$
Aggregate News	4.079***	1.097***	1.216***	1.088***	0.996***
	(21.30)	(17.26)	(18.92)	(17.75)	(18.22)
News-day Returns	13.111***	0.146***	0.011	-0.030	-0.009
	(16.43)	(3.34)	(0.50)	(-1.22)	(-0.53)
Daily	0.576***	0.018***	0.016***	0.011***	0.010***
	(208.71)	(37.61)	(47.52)	(41.73)	(44.81)
Weekly	0.566***	0.042***	0.027***	0.020***	0.019***
	(82.17)	(32.43)	(33.76)	(37.64)	(39.47)
Neutral News Included	0.380***	0.034***	0.042***	0.029***	0.020***
	(30.91)	(23.93)	(46.75)	(36.06)	(27.50)
Negative vs. Positive	0.258***	0.025***	0.029***	0.021***	0.014***
	(23.00)	(30.28)	(33.12)	(34.19)	(29.92)
Decile Portfolios	0.558***	0.043***	0.050***	0.034***	0.022***
	(26.69)	(26.29)	(47.77)	(33.02)	(22.19)

Table IA2: News and Firm Fundamentals

This table examines the relation between the news (*$News_t$*) and firm fundamentals (ROA, SUE) using pool regressions. The dependent variables are future ROAs in Panel A and future SUEs in Panel B respectively. ROA_{t+1} (SUE_{t+1}) shows the 1-quarter average ROA (SUE) of each portfolio in quarter t+1; ROA_{t+2} (SUE_{t+2}) shows the 1-quarter average ROA (SUE) in quarter t+2; $ROA_{t+3,t+4}$ ($SUE_{t+3,t+4}$) shows the average ROA (SUE) over 2 quarters from t+3 to t+4; and $ROA_{t+5,t+8}$ ($SUE_{t+5,t+8}$) shows the average ROA (SUE) over 4 quarters from t+5 to t+8. *Total Assets*$_t$ is the total assets at the end of previous year. B/M_t is the book-to-market ratio defined as the ratio of book value of equity to market value of equity in the previous year. *Asset Growth*$_t$ is the growth rate in total assets. *Leverage*$_t$ is the ratio of total long-term debt over total assets. The regressions also include the year-quarter fixed effect and industry fixed effect. The table also reports the number of observations (*NObs*), number of firms (*Firms*), and adjusted R square (*Adj-*R^2). The sample period is from January 2000 to October 2014. Standard errors are clustered by firms. The *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Panel A: Future ROA				
	ROA_{t+1}	ROA_{t+2}	$ROA_{t+3,\, t+4}$	$ROA_{t+5,\, t+8}$
Variables	M1	M2	M3	M4
$News_t$	**2.124*** **	**2.130*** **	**1.820*** **	**1.281*** **
	(13.19)	**(11.57)**	**(8.91)**	**(6.21)**
ROA_t	0.537***	0.499***	0.501***	0.443***
	(29.70)	(17.35)	(18.08)	(17.59)
Total Assets$_t$	0.423**	0.521**	0.537**	0.584**
	(2.21)	(2.23)	(2.17)	(2.03)
B/M_t	-0.239***	-0.215***	-0.155***	-0.068
	(-8.42)	(-6.21)	(-4.06)	(-1.61)
Asset Growth$_t$	-0.002**	-0.003**	-0.004***	-0.004***
	(-2.25)	(-2.00)	(-3.81)	(-5.22)
Leverage$_t$	0.565***	0.736***	0.972***	1.229***
	(5.02)	(5.45)	(6.69)	(6.86)
Year-Quarter fixed effect	Yes	Yes	Yes	Yes
Industry fixed effect	Yes	Yes	Yes	Yes
NObs	190,246	187,902	182,913	163,751
Firms	6450	6382	6193	5707
Adj-R^2	34.3%	29.0%	33.3%	31.7%

Table IA2: News and Firm Fundamentals (Continued)

Panel B: Future SUE				
	SUE_{t+1}	SUE_{t+2}	$SUE_{t+3,\ t+4}$	$SUE_{t+5,\ t+8}$
Variables	(1)	(2)	(3)	(4)
$News_t$	**3.464*****	**2.317*****	**1.719*****	**0.650****
	(11.85)	**(6.99)**	**(5.96)**	**(2.28)**
SUE_t	0.075***	0.070***	0.046***	0.060***
	(4.85)	(4.33)	(5.20)	(2.82)
$Total\ Assets_t$	1.252***	1.325***	1.276***	1.335***
	(3.87)	(3.82)	(3.45)	(3.08)
B/M_t	-0.627***	-0.600***	-0.400***	-0.229**
	(-5.51)	(-5.03)	(-4.07)	(-2.19)
$Asset\ Growth_t$	0.001***	0.001***	0.000	-0.012
	(4.31)	(2.61)	(0.30)	(-1.57)
$Leverage_t$	-0.045	-0.223	-0.085	-0.026
	(-0.25)	(-0.98)	(-0.44)	(-0.14)
Year-Quarter fixed effect	Yes	Yes	Yes	Yes
Industry fixed effect	Yes	Yes	Yes	Yes
NObs	116,233	113,916	102,694	83,602
Firms	4913	4830	4485	3722
Adj-R^2	1.9%	1.7%	2.5%	4.5%

Table IA3: Return Predictability of News—Robustness Tests

The table examines the robustness of return predictability of news by using different specifications. "*Aggregate News*" means that each month all stocks are grouped into five portfolios based on their aggregate news scores ($News_t$) which is the sum of ESS score of all news for a particular firm over a month t. "*News-day returns*" means that each month all stocks are grouped into five portfolios based on their News-day returns which is the average 3-day returns (-1, 1) around each news for all news for a particular firm over a month t. *"Daily"* and *"Weekly"* means that all stocks are grouped into five portfolios based on their news scores ($News_t$) at the end of day t and week t, respectively. *"Neutral News Included"* means that all stocks including those with zero news scores (neutral news or no news) are grouped into five portfolios based on their news scores ($News_t$) at the end of month t. *"Decile Portfolios"* means that all stocks are grouped into ten portfolios based on news scores ($News_t$) at the end of month t. For these four specifications, stocks in *Bad News* portfolio have the lowest news scores and stocks in *Good News* portfolio have the highest news scores. For the specification *"Negative vs. Positive"*, stocks in *Bad News* portfolio have the negative news scores and stocks in *Good News* portfolio have the positive news scores. "*Good-Bad*" is the hedge portfolio that is long in *Good News* portfolio and short in *Bad News* portfolio. We then compute the equally weighted next month (or week, or day) average return (R_{t+1}), CAPM alpha ($R_{CAPM,\,t+1}$), Fama and French three factor alpha ($R_{FF3,\,t+1}$), four factor alpha ($R_{FF4,\,t+1}$) and five factor alpha ($R_{FF5,\,t+1}$) for "Good-Bad" hedge portfolios. The sample period is from January 2000 to October 2014. Newey-West adjusted *t*-statistics are reported in the parentheses and ***, **, * denote 1%, 5%, and 10% significant levels, respectively.

Specifications	R_{t+1}	$R_{CAPM,\,t+1}$	$R_{FF3,\,t+1}$	$R_{FF4,\,t+1}$	$R_{FF5,\,t+1}$
Aggregate News	0.430**	0.498***	0.605***	0.547***	0.579***
	(2.50)	(2.74)	(3.49)	(3.53)	(3.20)
News-day Returns	0.666***	0.737***	0.700***	0.669***	0.729***
	(4.46)	(3.75)	(3.53)	(3.44)	(3.75)
Daily	0.531***	0.533***	0.532***	0.531***	0.527***
	(30.57)	(36.12)	(36.04)	(36.01)	(35.69)
Weekly	0.302***	0.313***	0.324***	0.305***	0.302***
	(7.44)	(8.71)	(9.02)	(8.81)	(8.44)
Neutral News Included	0.599***	0.674***	0.699***	0.657***	0.679***
	(4.32)	(4.91)	(4.99)	(5.10)	(4.68)
Negative vs. Positive	0.394***	0.456***	0.475***	0.441***	0.439***
	(3.34)	(3.98)	(4.11)	(4.14)	(3.69)
Decile Portfolios	0.877***	0.968***	0.965***	0.901***	0.947***
	(4.22)	(4.98)	(4.85)	(5.03)	(4.66)