Media-based Inter-Industry Network and Information Transmission

*

Li Guo[†]

Frank Weikai Li[‡]

This Draft: March 2019

Abstract

We construct a dynamic inter-industry network using a comprehensive sample of media news to examine how news travels across industries. Our analyses show that cross-industry news contains valuable information about firm fundamentals that is not fully captured by firms' own news or within-industry peers' news. Stock prices do not promptly incorporate cross-industry news, generating return predictability. Underraction to cross-industry news is more pronounced among smaller stocks that are more illiquid, more volatile, and have fewer analysts following. A long–short strategy exploiting cross-industry news yields annual alphas of over 10%.

Keywords: Media News; Information Diffusion; Inter-Industry Network
JEL Classification: G11; G12; G14

[*]We thank Qing Tong (discussant), Ximing Yin (discussant), Noolee Kim (discussant), Jun Tu, Johan Sulaeman, Roger Loh, Qifei Zhu, Chishen Wei, Hao Liang, and seminar participants at the 2017 SMU Finance Summer Camp and 2018 Singapore Scholars Symposium for helpful comments. All errors are our own.

[†]Singapore Management University, Lee Kong Chian School of Business (Email: liguo.2014@pbs.smu.edu.sg)

[‡]Singapore Management University, Lee Kong Chian School of Business (Email: wkli@smu.edu.sg)

1 Introduction

In modern economy, firms are intertwined through various observable and subtle economic links and these relations could easily transcend the traditional industry boundaries. Previous studies show that investors, largely due to their limited attention, underreact to the value-relevant information contained in the news of economically-related firms, which then leads to cross-firm return predictability.[1] While these prior studies mainly focus on customer-supplier linkage to understand inter-sectorial relationship[2], it may not fully capture the complex interdependence among firms operate in different sectors. In this paper, we use media news to construct a comprehensive inter-industry network and examine information transmission across industries.

There are several reasons to believe that media news may capture inter-industry relationship beyond what is captured by customer-supplier relation. First, media news, due to journalists' wisdom of crowd, is a comprehensive measure of cross-firm connections including product similarity, geographic overlap, business alliance, labor market competition, and regulatory impact. Second, unlike the customer-supplier relation which is updated infrequently, media news provide timely information about the dynamics of industry interdependence. Last but not the least, while prior studies mainly use realized stock returns as proxy for news, our use of media news is potentially less noisy and also contains soft information that may not be quickly impounded into prices. As a result, the media-based cross-industry network could complement the customer-supplier linkage in capturing the complexity of industry interdependence.

We construct the cross-industry network based on the number of media news simultaneously mentioning firms operating in two different industries. We then validate our

[1] See, for example, Cohen and Frazzini (2008); Menzly and Ozbas (2010); Hong, Torous and Valkanov (2007), and Rapach, Strauss, Tu and Zhou (2015); Lee, Sun, Wang and Zhang (2018).

[2] Related studies include Cohen and Frazzini (2008); Menzly and Ozbas (2010), Ahern (2013), Ahern and Harford (2014), Kelly, Lustig and Van Nieuwerburgh (2013), Herskovic, Kelly, Lustig and Van Nieuwerburgh (2016), Long Jr and Plosser (1983), Loualiche *et al.* (2014) and Oberfield (2012)

media-based industry network using the price delayless measure of Hou and Moskowitz (2005). Using techniques from graph theory, we calculate the centrality of each industry by measuring the strength of connections between this industry and all other industries in the economy. We find industries in the highest quintile of eigen-centrality (degree-centrality) have, on average, an average price delayness that is 6.58% (6.22%) lower than industries in the lowest quintile of eigen-centrality (degree centrality). This result is statistically significant and economically meaningful. We observe similar pattern when measuring the pairwise inter-industry connections. A pair of industries connected by the largest number of media news, on average, has an 8.15% shorter delay in incorporating the pair industry's news, compared to pair of industries connected by very few news. Next, we examine the pairwise correlation of investors' attention between the connected industry pairs, where attetion is measured by Google and Bloomberg search volumes. Our result shows that the industry pairs that are more frequently connected tend to have higher correlation of investors' attention. Investors' correlated search activities among connected industries suggest that media news draws investors' attention toward connected industries and facilitate the incorporation of cross-industry news into stock prices.

Having validate the media-based industry network, we next examine hows news diffuse across industries. Recent work suggests that media news contains soft information about firms' fundamentals, and has incremental predictive power for firms' future performance.[3] The literature, however, almost exclusively focus on the soft information contained in firms' own news. In this paper, we deviate from prior studies in examining the information contained in cross-industry news.

Specifically, we conduct textual analysis using the Thomson Reuters News Archive and construct news tones for each of the Fama-French 30 industry categories, where news tone is measured as the proportion of negative words following Tetlock *et al.* (2008).

[3] For example, Tetlock, Saar-Tsechansky and Macskassy (2008) found that negative words predict future earnings, and Bushee, Core and Hamm (2010) showed that the media serves as an information intermediary which incrementally contributes to firms' information environment.

We test the informativeness of cross-industry news by examining its predictability for firms' future unexpected earnings. If cross-industry news is incrementally useful, it should predict firms' fundamentals. Indeed, our analysis reveals the strong predictability of cross-industry news tone for firms' earnings news, even after controlling other predictor of firm fundamentals and firms' own news. The result also reveals the complexity of the cross-industry network in the real economy, as the coefficients in front of the cross-industry news tones exhibit substantial heterogeneity across industries.

Next, we link the cross-industry news to cross-industry return predictability. However, we do not test return predictability directly at the industry level. Instead, we estimate the value implication of cross-industry news for each firm and examine its return predictability at stock level. There are some reasons for doing this. First, even for firms within the same industry, they may react differently towards cross-industry news depending on their competitive positions within the industry. If this is the case, our approach would fully explore firms' heterogeneous exposure to the cross-industry information. Second, due to limited number of industries, industry-level test may lack the power to detect the informativeness of cross-industry news. Our stock-level test circumvent this power issue since we have on average 2,234 firms in each cross section, generating wide spread in terms of cross-industry news signal.

Return predictability test show that stock prices incorporate the information embedded in the cross-industry news with a significant delay. We obtain consistent results using both Fama-MacBeth regression and portfolio sorting. For example, a weekly-rebalanced, long-short portfolio that long stocks with positive CIS and short those with negative CIS generates Carhart (1997) four-factor alpha of more than 10% annually. The profitability of trading on cross-industry news survives after accounting for reasonable estimate of transaction costs. We further explore the horizon over which cross-indutry news diffuse into stock prices, and find news travel slowly in our case. The long-short portfolio based on cross-industry news still generates a sizeable alpha even 10 weeks after the news is announced to public. On the contrary, we find market prices impound firm-specific news relatively quickly, as the return to a long-short portfolio based on firms' own news fully dissipate after 4 weeks. In Fama-MacBeth regressions with firms' own news and news of within-industry peers as controls, we find cross-industry news continues to be a significant predictor of future returns, suggesting that cross-industry news contain novel information not captured by these alternative information sources. We also construct a cross-industry return signal for each stock and show that the cross-industry news signal explains the predictability of cross-industry return signal but not the other way around. This supports our use of media news as measure of cross-industry information, as it is less noisy than realized returns and also contain soft information.

We conduct several subsample tests based on firms' information environments, arbitrage frictions, and aggregate uncertainty. Our proxies for firms' information environments include firm size, analyst coverage and forecast dispersion. The results show that the return predictability of cross-industry news is much more pronounced among stocks with poor public information environments, such as small stocks with thin analyst coverage. In a similar vein, we find the return predictability of cross-industry news is larger for stocks that are more difficult to arbitrage, such as high volatility and illiquid stocks. In addition, cross-industry news seem to diffuse more slowly during highly uncertain periods, as proxied by higher VIX and more dispersed news signal.

The paper contributes to several strands of literature. First, our work relates to the empirical studies on gradual information diffusion among economically linked firms and industries. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) document return predictability from customer firms/sectors to supplier firms/sectors. Hong *et al.* (2007) show that the returns of the leading industry lead the market returns. Lee *et al.* (2018) show that returns of technology-linked firms have strong predictive power for focal firm returns. Parsons, Sabbatucci and Titman (2018) document lead-lag effects in stock returns between co-headquartered firms operating in different sectors. Our study is similar in spirit, but examines information diffusion along the inter-industry network extracted from media news. As we have conjectured, our media-based industry network has the advantage of being a more comprehensive measure of cross-firm connections and is also dynamic.

Second, this paper contributes to the growing literature on quantifying soft information in news and examine its value implications for firms' fundamentals and stock prices. Tetlock (2007) analyzes the content of a commentary section in the Wall Street Journal, and finds that pessimistic words predict lower stock returns the next day. Davis, Piger and Sedor (2006), Tetlock *et al.* (2008), and Demers and Vega (2011) extract the tone from firm-specific news and show its informativeness for firms' future earnings and stock returns. Our study builds upon this literature and shows that tones of cross-industry news contain valuable information about firm fundamentals beyond what is captured by firms' own news.

Third, this paper also enhances our understanding of the role of media as an information intermediary. Fang and Peress (2009) show firms with lower media coverage have higher expected returns, as predicted by Merton (1987) when investors have incomplete information and market is segmented. Peress (2014) uses newspaper strikes as an exogenous shock, and show that media affect the stock market by improving the speed of information diffusion among investors. Engelberg and Parsons (2011) document direct evidence of local media coverage affecting local investors' trading activities. Bushee *et al.* (2010) find that media coverage reduces information asymmetry around earnings announcements through broad dissemination of information. Our paper differs from these studies by showing media news help facilitate the information transimision across industries and firms.

The rest of the paper is organized as follows. Section 2 describes the data used in this paper, and explains how the Cross-Industry News Signal (CIS) is constructed. Section 3 constructs and validates the media-based inter-industry network. Section 4 presents results on the information diffusion of cross-industry news. Section 5 explores the channels through which cross-industry news diffuse into stock prices. The last section concludes.

2 Data and Methodology

2.1 Data and Variables

The data used in this paper is collected from five major datasets. Media news is from Thomson Reuters. Analysts' annual earnings forecasts and other related information are obtained from the I/B/E/S. The institutional fund flow data is collected from the EPRF database. The data for firms' fundamentals and stock market variables are obtained from the Compustat and CRSP databases, respectively.

We construct the news sample using firm-specific news articles for all U.S. public firms from January 1996 to December 2014. We require the news articles to be novel news, which means that it is the first release or record by Thomson Reuters. We classify news items into Fama-French 30 industry categories according to the firms' RIC mentioned in each news article.[4] In total, we retrieve 11.63 million news stories from the Reuters News Archive database.

To construct the media-based inter-industry network, we first convert news data every year into the matrix $\mathcal{M}_t$ below:

[4]The RIC is made up primarily of the security's ticker symbol, optionally followed by a period and exchange code based on the name of the stock exchange using that ticker. For instance, IBM.N is a valid RIC, referring to IBM being traded on the New York Stock Exchange. By extracting ticker symbol from RIC, we are able to match it with CRSP permno.

$$\mathcal{M}_t = \begin{array}{c} \\ \scriptstyle Industry_1 \\ \scriptstyle Industry_2 \\ \vdots \\ \scriptstyle Industry_N \end{array} \begin{array}{c} \begin{array}{cccc} \scriptstyle news_1 & \scriptstyle news_2 & \scriptstyle \cdots & \scriptstyle news_{K_t} \end{array} \\ \begin{bmatrix} Occr_{1,t}^{1} & Occr_{1,t}^{2} & \cdots & Occr_{1,t}^{K_t} \\ Occr_{2,t}^{1} & Occr_{2,t}^{2} & \cdots & Occr_{2,t}^{K_t} \\ \vdots & \vdots & \ddots & \vdots \\ Occr_{N,t}^{1} & Occr_{N,t}^{2} & \cdots & Occr_{N,t}^{K_t} \end{bmatrix} \end{array}, \tag{2.1}$$

where N is the total number of industries in the sample, K_t is the total number of news each year, and $Occr_{n,t}^{k}$ equals 1 if a stock in industry n is mentioned by a news article, k. Based on $\mathcal{M}_t$, we then obtain the *weighted adjacency matrix*, $\mathcal{W}_t$, that measures the strength of connectivities between two industries:

$$\mathcal{W}_t = \mathcal{M}_t\mathcal{M}_t^{\top} = \begin{array}{c} \\ \scriptstyle industry_1 \\ \scriptstyle industry_2 \\ \vdots \\ \scriptstyle industry_N \end{array} \begin{array}{c} \begin{array}{cccc} \scriptstyle industry_1 & \scriptstyle industry_2 & \scriptstyle \cdots & \scriptstyle industry_N \end{array} \\ \begin{bmatrix} w_{11,t} & w_{12,t} & \cdots & w_{1N,t} \\ w_{21,t} & w_{22,t} & \cdots & w_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1,t} & w_{N2,t} & \cdots & w_{NN,t} \end{bmatrix} \end{array}, \tag{2.2}$$

where $w_{ij,t} = \sum_{k=1}^{K_t} Occr_{i,t}^{k} Occr_{j,t}^{k}$ with $i, j = 1, 2, \cdots, N$. Intuitively, when $i = j$, $w_{ii,t}$ counts the number of news mentioning the industry i at time t, and when $i \neq j$, $w_{ij,t}$, it counts the number of news co-mentioning industry i and j at time t. We then calculate the Eigen-centrality for the above weighted adjacency matrix after setting the diagonal element to 0. Different from the adjacency matrix, the Eigen-centrality for the weighted adjacency matrix considers the strength of connections between nodes. Specifically, Eigen-centrality is defined as follows:

$$\mathcal{W}_t\mathbf{x}_t = \lambda_{\max}\mathbf{x}_t, \text{ for each } t = 1, 2, \cdots, T, \tag{2.3}$$

where $\mathbf{x}_t = (Ctry_{1,t}, Ctry_{2,t}, \cdots, Ctry_{N,t})'$ and $Ctry_{i,t}$ is the eigenvector centrality score of industry i at time t.[5] We also construct the degree-centrality measure by counting the number of news connecting industry i and all other industries, namely:

$$Degree - centrality_{i,t} = \frac{\sum_{j \neq i} w_{ij,t}}{\sum_{i=1}^{i=N} \sum_{j \neq i} w_{ij,t}}.$$

We conduct textual analysis to quantify the information content of each news article using the word list of Loughran and McDonald (2011). We use a variation of the approach in Hu and Liu (2004) to account for sentiment negation. If the word distance between a negation word ("not," "never," "no," "neither," "nor," "none," "n't") and the sentiment word is not larger than five, the positive or negative polarity of the word is changed to the opposite of its original polarity. Following Tetlock *et al.* (2008), we measure the tone of each news article using the negative word ratio according to the following equation:

$$\text{Tone} = \frac{\text{\# of negative words}}{\text{Total \# of words in the news article}}.$$

We then compute the firm-specific news tone by averaging the tone for all news articles related to the firm i at time t:

$$\text{Firm News}_{i,t} = \frac{\sum_{d=1}^{D} \text{Tone}_{i,d}}{\text{D}}.$$

where D is the total number of firm-specific news at time t. We define the news tone for firm i's industry peers as the average news tone of peer firms within the same industry as firm i:

$$\text{Peer News}_{i,t} = \frac{\sum_{k=1}^{K} \text{Firm News}_{k,t}}{\text{K}},$$

[5] According to Segarra and Ribeiro (2016), eigen-centrality shows stable property for the weighted adjacency matrix.

where i $\neq$ k and K is the total number of firms in industry excluding firm i. Similarly, we define the cross-industry news of firm i as the average firm-specific news tone of a cross-industry $\mathbf{J}$, namely:

$$\text{Cross-Industry News}_{i,\mathbf{J},t} = \frac{\sum_{j=1}^{J} \text{Firm News}_{j,t}}{J},$$

where $\mathbf{J} \in \{1, 2..., N-1\}$, N is the total number of industries, and J is the total number of firms in industry $\mathbf{J}$. To control for the effect of media coverage (Fang and Peress (2009)), we also calculate the number of firm-specific news, the number of industry peer news and the number of cross-industry news as additional controls.

We use the firm's standardized unexpected earnings (SUE) as a proxy for the firm fundamentals. Following Bernard and Thomas (1989), SUE is defined as:

$$\text{UE}_t = \text{E}_t - \text{E}_{t-4}$$

$$\text{SUE}_t = \frac{\text{UE}_\text{t} - \overline{\text{UE}}_\text{t}}{\text{Std}(\text{UE}_t)},$$

where E_t is the firm's earnings in quarter t. $\overline{\text{UE}}$ and Std(UE) are the mean and volatility of the unexpected earnings calculated using firm's previous 20 quarters of unexpected earnings, respectively. We also include control variables including firm size, book-to-market ratio (B/M), turnover, three measures of recent stock performance, and analyst forecast dispersion. Firm size (log(market capitalization)) and B/M are calculated at the end of the preceding calendar year, following Fama and French (1993). The turnover is the natural log of number of shares traded divided by shares outstanding (Log(Share Turnover)) at the end of the preceding calendar year. We calculate analyst dispersion as the standard deviation of analysts' earnings forecasts 3 to 30 days prior to the earnings announcement scaled by earnings volatility.

Following Tetlock *et al.* (2008), we calculate past returns based on a simple event study methodology. We chose the analysts' forecast announcement day or earnings announcement day as the event day in accordance with the dependent variable. Specifically, we calculate the expected return using the Fama-French three-factor model with an estimation window of [-252,-31] trading days before the event day t. We also calculate the abnormal return on day $t-2$ before the event day CAR_{t-2}, and the cumulative abnormal return in the [-30,-3] trading day window before the event day, denoted as $\text{CAR}_{t-30,t-3}$.

Following Druz *et al.* (2015), we include some firm characteristics as control variables. Market return is defined as the percent value-weighted market return for the period starting 5 days after an earnings announcement for the quarter t-1 and ending 5 days before the earnings announcement for the quarter t. Momentum is defined as the firm's buy-and-hold return over the previous 6 months. Illiquidity is defined as the absolute value of the stock return scaled by the dollar trading volume. Leverage is defined as the long-term debt scaled by the sum of the long-term debt and equity market capitalization. Institutional Ownership is defined as the number of shares held by 13F institutions scaled by shares outstanding. Return volatility is the standard deviation of the monthly return over the previous 48 months. In some specifications, we also include firm fixed effects and year fixed effects.

Panel C of Table 1 presents the summary statistics for the variables related to media news. For an average firm in our sample, there are 57,507 cross-industry news, 1,237 industry peer news, and 28 firm-specific news within 90 days before the earnings announcement. The # of Cross-Industry News is much larger than the number of firm-specific news, suggesting that more information is potentially revealed by cross-industry news. The mean industry-level news tone is 0.045, ranging from 0.038 to 0.052. As expected, the volatility of industry-level news tone is much smaller than firm-specific news tone. Following Tetlock *et al.* (2008), we standardize all news tones to make it comparable across industries.

< Insert Table 1 here >

2.2 Cross-Industry News Signal

To examine the information content of cross-industry news, one needs to measure the overall impact of cross-industry news for each individual firm. This is challenging due to the complex inter-industry relationship, and the same industry's news may have differential value implication for different firms. In this paper, we use a machine learning approach to extract the information from multiple cross-industry news, which we denote as the cross-industry news signal (CIS).

The approach consists of three steps. First, we calculate the news tone of each industry J over the most recent week $t-1$, denoted as Cross-Industry News$_{i,\mathbf{J},t-1}$. Next, we predict stock i's week-t return using the news tone of firm i's cross industries over the week $t-1$. The predictive regression is estimated as follows:

$$\mathrm{r}_{i,t} = \alpha_i + \sum_{\mathbf{J}=1}^{N-1} b_{i,\mathbf{J},t} \text{Cross-Industry News}_{\mathrm{i},\mathbf{J},\mathrm{t}-1} + \epsilon_{i,t}, \text{for t = 1,..., T,} \tag{2.4}$$

where $r_{i,t}$ is the week-t return of stock i in excess of the risk-free return, and N is the total number of industries. We require at least 260 weekly observations for each firm, and set the initial estimation window at 208 weeks (4 years of observation).

Moreover, to improve out-of-sample prediction and avoid model overfitting, we use the adaptive LASSO method following Zou (2006). The adaptive lasso includes parameter weights in the LASSO penalty term to achieve the oracle properties for appropriate weights. The adaptive LASSO estimates are defined as:

$$\hat{b}_i^* = \mathrm{argmin} \| r_{i,t} - \alpha_i - \sum_{\mathbf{J}=1}^{N-1} b_{i,\mathbf{J}} \text{Cross-Industry News}_{\mathrm{i},\mathbf{J},\mathrm{t}-1} \|^2 + \lambda_i \sum_{\mathbf{J}=1}^{N-1} \hat{w}_i |b_{i,\mathbf{J}}|,$$

where Cross-Industry News$_{i,\mathbf{J},t-1}$ is the standardized news tone of cross-industry $\mathbf{J}$, $\hat{b}_i^* = (\hat{b}_{i,1}^*, ..., \hat{b}_{i,N-1}^*)'$ is the N-1 vector of adaptive LASSO estimates, λ_i is a nonnegative regularization parameter, and $\hat{w}_{i,\mathbf{J}}$ is the weight assigned to $|b_{i,\mathbf{J}}|$ in the penalty term. The adaptive LASSO uses the L1-norm penalty to prevent overfitting, and enables the selection of the most informative predictors.

Using the estimated $\hat{b}_i^*$, we then predict the out-of-sample return in week $t+1$ using the cross-industry news available at time t and denote it as CIS:

$$\text{CIS}_{i,t} = \alpha_i + \sum_{\mathbf{J}=1}^{N-1} E_t[b_{i,\mathbf{J},t+1}]\text{Cross-Industry News}_{\mathrm{i},\mathbf{J},\mathrm{t}},$$

where $E_t[b_{i,\mathbf{J},t+1}]$ is the estimated coefficient from equation 2.4 and is defined as $E_t[b_{i,\mathbf{J},t+1}] = b_{i,\mathbf{J},t}$. The cross-industry news signal is a real-time predictor of stock returns and does not suffer from look-forward bias.[6]

3 Media-based Inter-industry Network

3.1 Properties of the Media-based Inter-industry Network

In this section, we first show that our media-based inter-industry network reveals complex and dynamic inter-industry relationship, beyond what is captured by the customer-supplier linkages. Figure 1 illustrates the time-varying inter-industry network for the Fama-French 30 industries based on Thomson Reuters News data from 1996 to 2014. We define two industries as connected if a piece of news article simultaneously mentions stocks in the two industries. The thickness of an edge reflects the degree of inter-industry connections, as determined by the number of news connecting an industry pair. The node size denotes the eigenvector centrality of an industry. The figure shows two stylized facts. First, our media-based inter-industry network varies significantly over time, suggesting a dynamic industry interdependence. Moreover, we observe that the inter-industry connections become stronger in the recent years. Second, unlike customer-supplier relation,

[6] In the empirical analysis below, we only use stocks with negative CIS due to the uninformativeness of positive CIS.

Manufacturing-related industries are usually not the central industries in our media-based industry network. Instead, the more central industries (represented by the larger node size) seem to be Business Equipment, Personal and Business Services, and Finance industry. Overall, the analysis suggests that our media-based inter-industry network is dynamic and distinct from the traditional customer-supplier network, and is potentially a more comprehensive measure of inter-industry relationship.

< Insert Figure 1 here >

3.2 Media-based Inter-industry Network and Price Delayness

To verify that our media-based industry network indeed captures the network position of industries, we first link industry-level centrality based on the media network with the price delayness measure of Hou and Moskowitz (2005). The idea is that more central industries should more quickly incorporate economy-wide shocks. Panel A of Table 1 reports the summary statistics of cross industry media connection, including eigenvector centrality, degree centrality, the frequency that an industry is assigned into a low (high) eigenvector centrality and low (high) degree centrality and the delayness measure. Consistent with Figure 1, Financials, Business Equipment, Services and Retails are the most important nodes in the media network, while Coal, Oil and Mines tend to be the periphery groups. Moreover, the industries with the highest network centrality tend to have a lower delayness measure.

In Table 2, we sort all industries into five quintiles based on the eigen-centrality (degree-centrality) and reports the average delayness measure for each group over 1996 to 2014. Column 1 of Table 2 shows the average delayness for quintiles sorted on eigen-centrality. The industries with the lowest eigen-centrality has an average price delayness of 7.90%, compared to 1.32% for industries in the highest quintile. The differences in price delayness between the highest and lowest quintiles of industries is -6.58% and highly significant. We observe similar pattern between industries' network centrality and price delayness using the degree-centrality measure in Column 2 of Table 2.

< Insert Table 2 here >

In Table 3, we report the cross-industry information delayness for five groups sorted on the number of news mentioning the given industry pair. Our approach to construct cross-industry delayness measure is similar to Hou and Moskowitz (2005). For an industry pair A and B, the delayness of industry B's news on industry A's return, $Delay_{B\rightarrow A}$, is the fraction of industry A's returns explained by industry B's lagged returns. More specifically, the measure is one minus the ratio of the R^2 from regression (4.1) by restricting $\delta_j^{-n} = 0, n \in [1, 4]$, over the R^2 from regression (4.1) without restrictions.

$$\mathrm{r}_{A,t} = \alpha_j + \beta_A r_{B,t} + \sum_{n=1}^{4} \delta_A^{-n} r_{B,t-n} + \epsilon_{A,t}, \qquad (4.1)$$

where $r_{A,t}$ is the daily return of industry A and $r_{B,t}$ is the daily return of industry B. The pairwise information delayness between industry A and B is calculated as the average of $Delay_{A\rightarrow B}$ and $Delay_{B\rightarrow A}$. We then sort all industry pairs into five quintiles according to the # of news mentioning the paired industries and report the average pairwise delayness measure for each quintile. Consistent with our measure capturing cross-industry relationship, we find industry pairs that are more closely connected through media have much lower cross-industry price delayness. For example, the average delayness for industry pairs with the lowest media connection is 13.21%, more than twice the average delayness of the industry pairs with the strongest media-based connection. Overall, our validation test based on price delayness measure strongly support the notion that media is an important information intermediary that contribute to cross-industry information diffusion.

< Insert Table 3 here >

3.3 Media-based Inter-industry Network and Investor Attention

We further verify that investor attention is the underlying channel that contributes to the reduced cross-industry information delay. To test, we use both Google and Bloomberg search volume index as direct proxy for investor attention. Following Da *et al.* (2011), we convert the raw Goolge search volume into abnormal search volume (ASV) defined as follows:

$$ASV_t = \frac{SVI_t}{\text{average SVI from } weekday_{t-260} \text{ to } weekday_{t-21}} - 1$$

Following Ben-Rephael *et al.* (2017), we convert Bloomberg's numerical search scores into continuous values, using the conditional means of truncated normal distribution. Under the normal distributional assumption, the corresponding values are -0.350, 1.045, 1.409, 1.647, and 2.154. We sort industry pairs into 5 quintiles according to the number of news connecting the industry pair and calculate the correlation of search volume within each industry pair. Figure 2 shows that the investors' co-search activities for the industry pairs through Google and Bloomberg monotonically increase with the number of news connecting the industry pair. The average correlations for industry pair with the highest number of news connecting the industry pair are 36.67% and 61.10% for Google search and Bloomberg search, respectively. The average correlations for industry pairs with the lowest number of connecting news are only 9.76% and 23.28% for Google search and Bloomberg search, respectively. These results suggest that news mentioning industry pair draws investors' attention towards the connected industries, and faciliate the information transmission across industry pairs.

< Insert Figure 2 here >

4 Cross-Industry News and Information Transmission

4.1 Cross-Industry News and Firm Fundamentals

In this section, we examine whether cross-industry news contain value-relevant information about firm fundamentals. We perform the following regression analysis:

$$\text{SUE}_{it} = \alpha_i + \sum_{\mathbf{J}=1}^{N-1} \beta_{\mathbf{J}} \text{Cross Industry News}_{i,\mathbf{J},t-90,t-3} + \gamma^{`}\text{X} + \epsilon_{it},$$

The dependent variable, SUE, is firms' standardized unexpected earnings following Bernard and Thomas (1989). Cross-Industry News$_{i,\mathbf{J},t-90,t-3}$ is the news tone of industries J over the period (t-90, t-3) relative to the earnings announcement day t. The control variables include firm-specific news tone, news tones of within-industry peer firms, # of firm-specific news, # of news of within-industry peer firms, # of cross-industry news. We also include those controls suggested by Tetlock *et al.* (2008), including firms' lagged earnings (proxied by last quarter's SUE, lagSUE), Size, B/M, Turnover, three measures of recent stock returns (AR$_{t-252,t-31}$, CAR$_{t-30,t-3}$ and AR$_{t-2}$), analysts' earnings forecast revisions (Forecast Revision), and analyst forecast dispersion (Analyst Dispersion). Besides, we further control other variables documented in prior literatures (Jegadeesh, Kim, Krische and Lee (2004) and Druz, Wagner and Zeckhauser (2015), among others), including a dummy variable indicating news coverage ($I_{newscoverage}$), Consensus Forecast, Management Forecast, Earnings Surprise, Return Volatility, Market Return, Institutional Ownership, Leverage, Momentum, Illiquidity and Overconfidence.

< Insert Table 4 here >

Table 4 presents the panel regression results, with standard errors clustered at firm level. In Panel A, we only include the news tone of one cross industry in the regression. The first three columns show the estimated coefficients, T-value and adjusted R^2 for the univariate regression. The middle three columns report the corresponding results that follow the specification of Tetlock *et al.* (2008). In the last three columns, we added all control variables. The results are consistent across different specifications. We find most cross-industry news negatively predict individual firm's earnings surprise. Only the news of the Coal industry positively predicted SUE. This is intuitive since the Coal industry serves as the most important raw inputs to other industries, thus a negative shock to the Coal industry causes a reduction of the input cost and positively affects the earnings of other downstream industries. In Panel B, we run the regression by including the news of all the cross-industries into one regression. The results change a lot due to the interactions of cross-industry news information. Indeed, some cross industry news become insignificant or even change their prediction signs. A number of industries remain strong predictors of individual firms' earnings, such as Food, Beer, Smoke, Books, Hlth, ElcEq, Autos, Mines, Paper and Trans.

On top of that, the loading on those industry news tones exhibits substantially positive and negative predictions on SUE, suggesting complex industry interdependencies that have bullish implications for some industries and bearish implications for others. This again emphasize the complexity of network effect in the real word. In this case, media news provides an new angle to understand the information diffusion across industries.

4.2 Cross-Industry News and Stock Returns

Having established that cross-industry news can predict firms' fundamentals, we examine if cross-industry news also provides novel information not fully reflected in stock prices. To test this, we examine the return predictability of CIS at stock level by running Fama-MacBeth regression. The advantage of Fama-MacBeth methodology is that one can control for a large set of firm characteristics that commonly associated with stock returns, including Lagged Return, Size, B/M, Leverage, Turnover, Return Volatility, Firm News, Industry News, # of Firm News, # of Industry news and # of Cross-Industry News.

Table 5 reports the time-series averages of the coefficients of the independent variables, and the t-statistics are Newey-West adjusted. The first three columns report the results using the whole sample period from 2000 to 2014 (year 1996 to 1999 is used as initial estimation window). The middle three columns report results for 2000 to 2007, and the last three columns show results for 2008 to 2014. Overall, CIS exerts a strong cross-sectional return predictability, and the results are robust across different specifications and sub-periods. The economic magnitude is also quite large. For the whole sample period, a one-standard-deviation increase in CIS increases the stock returns by 2.25%.

< **Insert Table 5 here** >

The significant coefficient in front of CIS in Fama-Macbeth regression suggests that a long-short strategy based on CIS should earn positive abnormal returns. At the end of each week, we sort all stocks with negative CIS into deciles and form an equal-weight long-short portfolio by shorting the stocks with most negative CIS and longing the stocks with the least negative CIS. We then hold the portfolio for one week and rebalance the portfolio at the end of each week. Figure 3 plots the cumulative return of the CIS long-short portfolio and the cumulative returns of the portfolio with all stocks included. The CIS long-short portfolio performs extremely well compared with the equal-weighted portfolio, suggesting the usefulness of the cross-industry news in predicting future price movements.

< Insert Figure 3 here >

Table 6 shows the weekly alphas of the long-short strategy based on CIS. Column (1), (4) and (7) report the CAPM adjusted alpha, column (2), (5) and (8) for the Fama-French three factor adjusted alpha, and column (3), (6) and (9) for the Carhart (1997) four-factor adjusted alpha. Standard errors are computed using the White (1980) heteroskedasticity-consistent covariance matrix. Consistent with the results from Fama-MacBeth regression, the CIS-based long-short strategy generates highly significant risk-adjusted returns of 20 bps per week, or around 10% annualized returns. In addition, the returns to CIS strategy have low exposures to common factors and are stable across sub-periods.

< Insert Table 6 here >

Given that the CIS-based strategy has relatively high portfolio turnover, we estimate the impact of transaction costs on its profitability. To that end, we re-calculate the returns and alphas to the CIS strategy under the assumption that a trader must incur a round-trip transaction cost between 1 and 10 bps. Table 7 reports the (annualized) raw and abnorml returns to CIS-based strategy under various trading cost assumptions. As a benchmark, we also show the returns to the long-short strategy based on firms' own news. The result shows that the CIS-based strategy survives after accounting for transaction cost, while the long-short strategy based on firms' own news is no longer profitable after accounting for reasonable level of transaction costs.

< Insert Table 7 here >

The above analysis suggests that cross-industry news contains valuable information about firms' future fundamentals. However, market seems to underreact to the information embedded in cross-industry news, leading to predictable returns. To further examine how cross-industry news slowly diffuse into stock price, we form long-short portfolios by skipping a period following the CIS signal. Specifically, we form the long-short portfolio at the end of each week based on the CIS signals from 2 to 10 weeks ago and hold the portfolio for 1 week. As a benchmark, we also report the returns to the long-short strategy based on firms' own news. Table 8 shows that cross-industry news diffuse more slowly into stock prices compared with firms' own news. The alphas of CIS-based strategy remains significant with 10.9% annualized return even 10 weeks after the signal. In sharp contrast, the long-short strategy based on firms' own news is no longer profitable 6 weeks after the signal. The result suggests that news travels slowly across industries.

< Insert Table 8 here >

4.3 Is Cross-industry News Explained by Alternative Information?

In this section, we examine the alternative explanation that cross-industry news may be explained by other sources of value-relevant information, including firms' own news, news from within-industry peers, and return-based cross-industry news. To investigate this possibility, we add the returns of three additional long-short portfolios in the time series regression, and the result is reported in Table 9. Column (1), (4) and (7) reports the alphas of CIS-based strategy after adding the portfolio returns based on the news of within-industry peer. Column (2), (5) and (8) reports the alphas after adding portfolio returns based on firms' own news. Columns (3), (6) and (9) reports the alphas after including portfolio returns based on cross-industry returns.

Indeed, the alphas to the CIS-based strategy reduce by around 1/3 after accounting for these alternative information sources, but remains positive and significant at 1% level. The result suggests that cross-industry news are partially overlapped with but not fully captured by these alternative sources of information, especially firms' own news and cross-industry returns.[7]

< Insert Table 9 here >

4.4 Impact of Economic Uncertainty

It is reasonable to expect that in more uncertain times, it takes longer time for investors to understand the implication of cross-industry news. To test this, we divide the whole sample into low and high uncertainty period based on the median value of economic uncertainty. Our proxies for economic uncertainty include VIX, economic policy uncertainty (Baker, Bloom and Davis (2016)), and a measure of market-wide news dispersion (Dzieliński and Hasseltoft (2015)). Market-wide news dispersion is defined as the cross-sectional standard deviation of news tone across firms. We then examine the profitability of the CIS strategy over the high and low uncertainty periods separately.

Table 10 reports the returns and alphas to the CIS-based strategy. The results are broadly consistent with our conjecture that the return predictability of CIS is indeed more pronounced in more uncertain market environment. For example, the annualized alpha of CIS strategy in high VIX periods is 4-5% higher than that in low VIX periods. We observe similar pattern using market-wide news dispersion, but not for economic policy uncertainty.

< Insert Table 10 here >

[7] In untabulated analysis, we show that cross-industry return strategy can be fully explained by the CIS-based strategy, suggesting that cross-industry news contain soft information not fully captured by cross-industry returns.

4.5 Cross-Sectional Heterogeneity in Information Environment and Arbitrage Frictions

Cross-industry news should be more valuable to firms with an opaque public information environment and limited firm-specific information. To test this, we use firm size, analyst coverage and analyst forecast dispersion as measures of a firm's information environment. We then examine the profitability of CIS strategy among firms with good and poor information environment, and plot the cumulative returns of the long-short portfolio in Figure 4. Consistent with our expectations, the CIS strategy performs much better among small firms with low analyst coverage, suggesting that cross-industry news travel slowly among such firms.

In addition to the firms' information environment, we also consider how the return predictability of CIS varies across our sample with differential degrees of arbitrage costs. Since cross-industry news is public information, the return predictability result suggests that sophisticated investors also fail to incorporate the information embedded in CIS and bring stock prices to full-information value. We thus expect our results to be more pronounced among stocks subject to greater limits to arbitrage. We use Amihud (2002) illiquidity and volatility as standard proxies of arbitrage costs. Graph D and E of Figure 4 plot the cumulative returns of the long-short CIS portfolio among subsamples with low and high arbitrage costs. The results are consistent with the limits-to-arbitrage predictions that CIS strategy generates much higher returns among more illiquid and volatile stocks.

< Insert Figure 4 here >

5 Channels of Cross-Industry Information Diffusion

Our evidence on market underreaction to cross-indutry news raises the question of how does the cross-industry news eventually get conveyed into stock price? In this section, we propose and test two underlying channels. First, the activites of information intermediary such as security analysts may facilitate the cross-industry information to be incorporated into stock prices. Second, cross-industry news may lead to subsequent firm-specific news being produced by media.

5.1 Cross-Industry News and Analyst Forecast Behavior

Analysts are widely recognized as an important information intermediary in the stock market. If cross-industry news contains valuable signals about firm fundamentals, analysts should incorporate this information into their earnings forecasts. To test this channel, we examine whether cross-industry news affects analysts' forecast behavior including forecast revisions and improvement in forecast accuracy. Specifically, we conduct the following regression:

$$\mathrm{Y}_{ijt} \quad = \quad \alpha + \beta_1 \text{Average Cross-industry News Tone}_{t-90,t-3} + \gamma^{`}\mathrm{X} + \epsilon_{ijt},$$

where Y_{ijt} is the forecast revision or improvement in forecast accuracy. Forecast revision is defined as the absolute change of two adjacent forecasts scaled by stock price at the end of the previous year. Forecast accuracy improvement is defined as the change in accuracy in two adjacent forecasts, where forecast accuracy is defined as the negative absolute value of the difference between actual earnings and forecasted earnings. X includes a set of explanatory variables, defined previously.

Table 11 presents the regression results. The first three columns report results for the forecast revision, and the last three columns show the results for forecast accuracy improvement. Consistent with our expectation, analysts tend to adjust their forecasts in the direction consistent with cross-industry news tone, and the revised forecast tends to be more accurate. Overall, the result suggest that analysts seem to incorporate cross-industry news into their earnings forecasts which facilitate price adjustment.

< Insert Table 11 here >

5.2 Cross-Industry News Leading Firm-Specific News

An independent channel through which cross-industry news transmit into stock prices may be through cross-industry news leading firm-specific news. To test this hypothesis, we run the Fama-MacBeth regression of firm-specific news tones on lagged CIS, controlling for lagged stock returns and other firm characteristics.

< Insert Table 12 here >

Table 12 reports the results. The first three columns show the results for the whole sample period, the middle three columns present results for 2000 to 2007, and the last three columns show the results for 2008 to 2014. Overall, we find that CIS consistently predicts next week's firm-specific news tone, even after controlling for lagged stock return. In terms of economic significance, a 1% increase in CIS leads to more negative firm-specific news tone by 7.54%. This evidence supports our hypothesis that cross-industry news provide leading clues about firms' fundamentals that is revealed eventually through firms' own news.

6 Conclusion

In this paper, we construct a dynamic inter-industry network using a comprehensive sample of media news to examine how news travels across industries. Our analyses show that cross-industry news contains valuable information about firm fundamentals that is not fully captured by firms' own news or within-industry peers' news. Stock prices do not promptly incorporate cross-industry news, generating return predictability. Underraction to cross-industry news is more pronounced among smaller stocks that are more illiquid, more volatile, and have fewer analysts following. A long–short strategy exploiting cross-industry news yields annual alphas of over 10%.

References

AHERN, K. R. (2013). Network centrality and the cross section of stock returns.

— and HARFORD, J. (2014). The importance of industry links in merger waves. *The Journal of Finance*, **69** (2), 527–576.

AMIHUD, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, **5** (1), 31–56.

BAKER, S. R., BLOOM, N. and DAVIS, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, **131** (4).

BEN-REPHAEL, A., DA, Z. and ISRAELSEN, R. D. (2017). It depends on where you search: institutional investor attention and underreaction to news. *The Review of Financial Studies*, **30** (9), 3009–3047.

BERNARD, V. L. and THOMAS, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting research*, pp. 1–36.

BUSHEE, B. J., CORE, G. W., J. E. and HAMM, S. J. (2010). The role of the business press as an information intermediary. *Journal of Accounting Research*, **48** (1), 1–19.

CARHART, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, **52** (1), 57–82.

COHEN, L. and FRAZZINI, A. (2008). Economic links and predictable returns. *The Journal of Finance*, **63** (4), 1977–2011.

DA, Z., ENGELBERG, J. and GAO, P. (2011). In search of attention. *The Journal of Finance*, **66** (5), 1461–1499.

DAVIS, A. K., PIGER, J. M. and SEDOR, L. M. (2006). Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases. *Working Paper*.

DEMERS, E. and VEGA, C. (2011). Linguistic tone in earnings press releases: News or noise? *Working Paper*.

DRUZ, M., WAGNER, A. F. and ZECKHAUSER, R. J. (2015). Tips and tells from managers: How analysts and the market read between the lines of conference calls. *NBER Working Paper*.

DZIELIŃSKI, M. and HASSELTOFT, H. (2015). News dispersion and investor disagreement. *Working Paper*.

ENGELBERG, J. E. and PARSONS, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, **66** (1), 67–97.

FAMA, E. F. and FRENCH, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, **33**, 3–56.

FANG, L. and PERESS, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, **64** (5), 2023–2052.

HERSKOVIC, B., KELLY, B., LUSTIG, H. and VAN NIEUWERBURGH, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics*, **119** (2), 249–283.

HONG, H., TOROUS, W. and VALKANOV, R. (2007). Do industries lead stock markets? *Journal of Financial Economics*, **83**, 367–396.

HOU, K. and MOSKOWITZ, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *The Review of Financial Studies*, **18** (3), 981–1020.

HU, M. and LIU, B. (2004). Mining and summarizing customer reviews. *ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177.

JEGADEESH, N., KIM, J., KRISCHE, S. D. and LEE, C. (2004). Analyzing the analysts: When do recommendations add value? *Journal of Finance*, **59** (3), 1083–1124.

KELLY, B., LUSTIG, H. and VAN NIEUWERBURGH, S. (2013). *Firm volatility in granular networks*. Tech. rep., National Bureau of Economic Research.

LEE, C., SUN, S. T., WANG, R. and ZHANG, R. (2018). Technological links and predictable returns. *Journal of Financial Economics, forthcoming*.

LONG JR, J. B. and PLOSSER, C. I. (1983). Real business cycles. *Journal of political Economy*, **91** (1), 39–69.

LOUALICHE, E. *et al.* (2014). Asset pricing with entry and imperfect competition. *Unpublished manuscript, Sloan School of Management, MIT*.

LOUGHRAN, T. and MCDONALD, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10ks. *Journal of Finance*, **66** (1), 35–65.

MENZLY, L. and OZBAS, O. (2010). Market segmentation and cross-predictability of returns. *The Journal of Finance*, **65** (4), 1555–1580.

MERTON, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, **42**, 483–510.

OBERFIELD, E. (2012). Business networks, production chains, and productivity: A theory of input-output architecture.

PERESS, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *The Journal of Finance*, **69** (5), 2007–2043.

RAPACH, D., STRAUSS, J., TU, J. and ZHOU, G. (2015). Industry interdependencies and cross-industry return predictability. *Working Paper*.

Segarra, S. and Ribeiro, A. (2016). Stability and continuity of centrality measures in weighted graphs. *IEEE Transactions on Signal Processing*, **64** (3), 543–555.

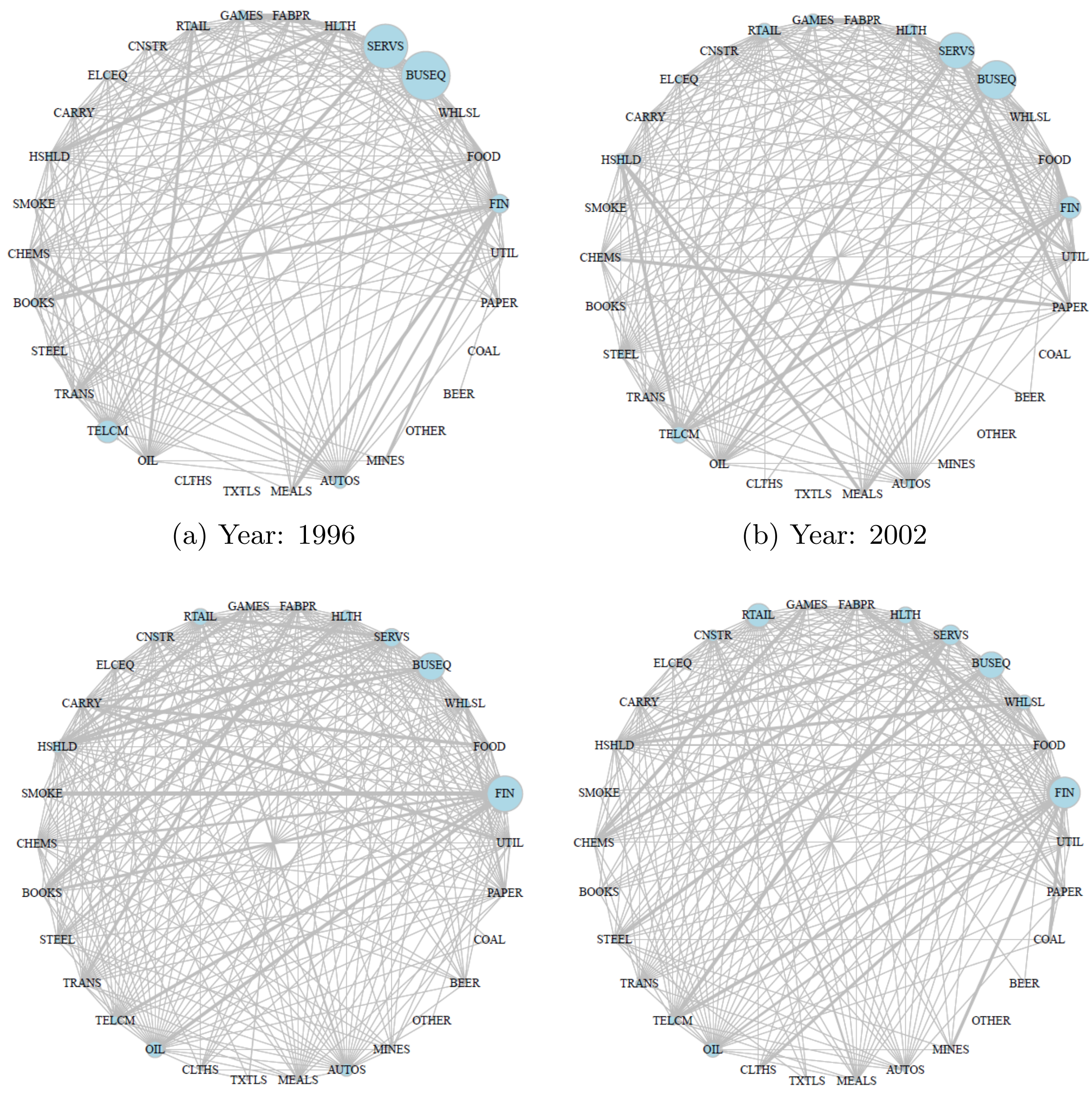
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, **62** (3), 1139–1168.

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

Wang, Y., Zhang, B. and Zhu, X. (2016). The momentum of news. *Working Paper*.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Journal of the Econometric Society*, pp. 817–838.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, **101** (476), 1418–1429.



(a) Year: 1996

(b) Year: 2002

(c) Year: 2008

(d) Year: 2014

Figure 1: **Media-based Inter-Industry Network.** This figure plots the media-based inter-industry network for Fama-French 30 industries in selective years. Two industries are connected if any news article simutaneously mentions stocks in these two industries. The thickness of an edge reflects the degree of connections between two industries, as measured by the number of news connecting two industries. The node size denotes the eigenvector centrality of an industry.

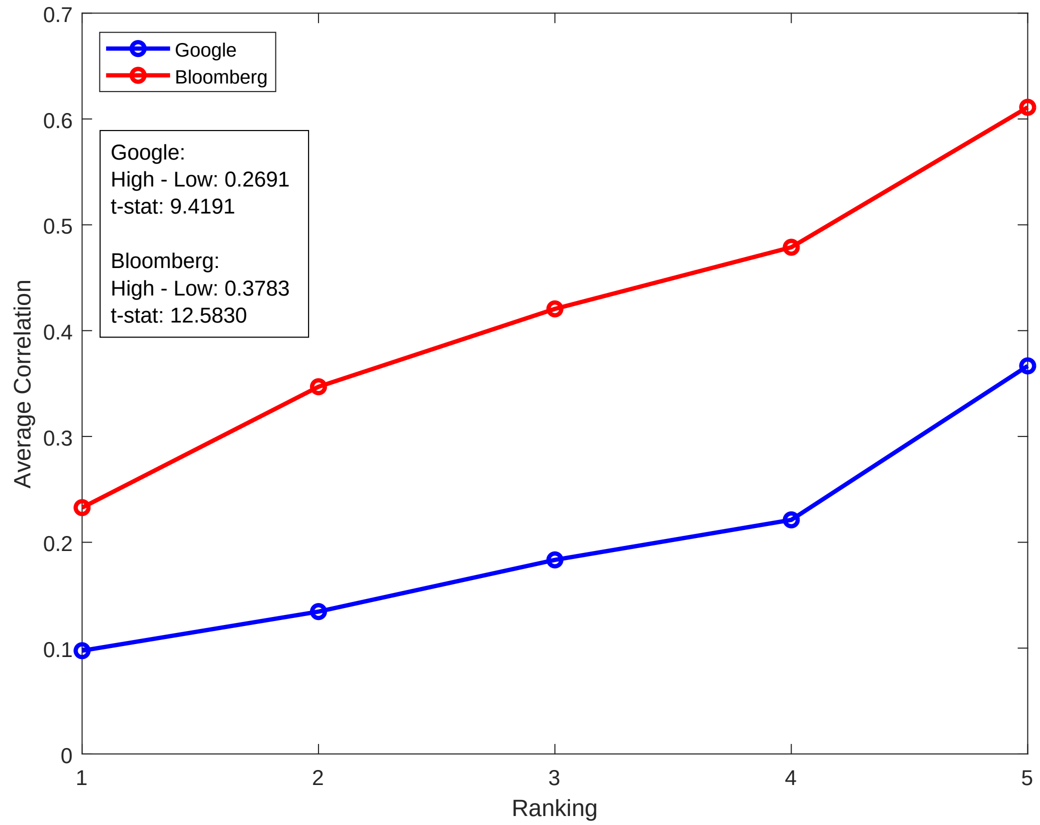


Figure 2: **Media-based Inter-Industry Connection and Investor Attention** This figure plots the investors' co-search activities through Google and Bloomberg for industry pairs. We sort industry pairs within Fama-French 30 industries into 5 quintiles based on the number of news connecting the industry pair and calculate the correlation of search volumes within each industry pair. The sample is from 1996 to 2014.

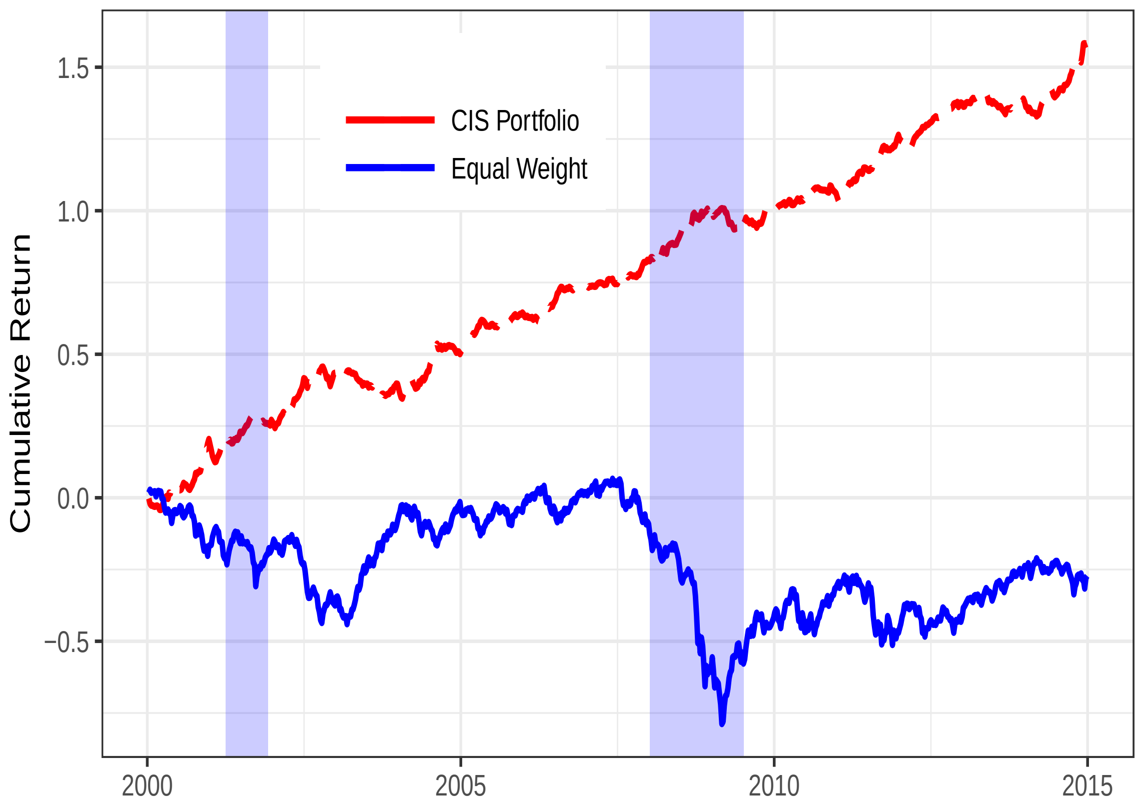


Figure 3: **Cumulative Returns of Long-Short Portfolio based on Cross-Industry News Signal (CIS).** This figure plots the cumulative returns of the long-short CIS portfolio (red line) and the portfolio holding all stocks with CIS available (blue line). At the end of each week, we sort all stocks with negative CIS into deciles and form an equal-weighted long-short portfolio by shorting the stocks with the most negative CIS and longing the stocks with the least negative CIS. We then hold the portfolio for 1 week and rebalance at the end of each week. The sample period runs from January 2000 to December 2014.

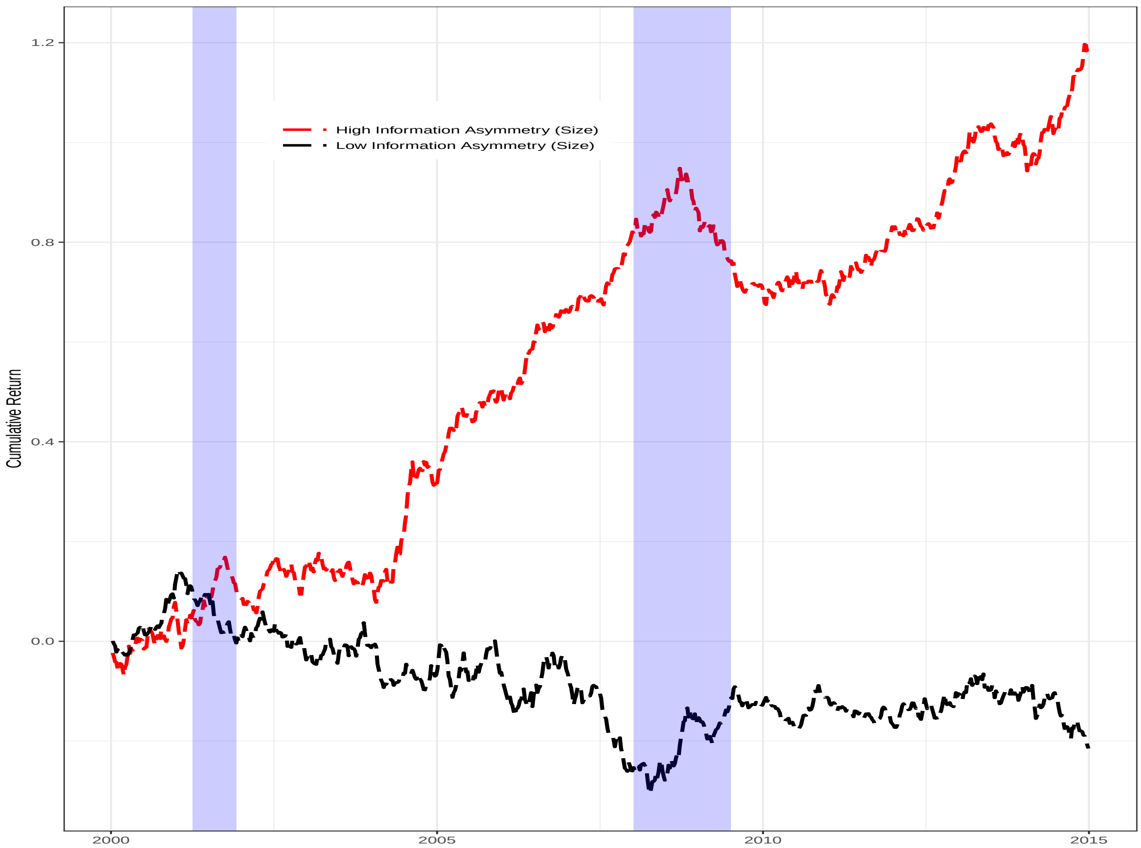
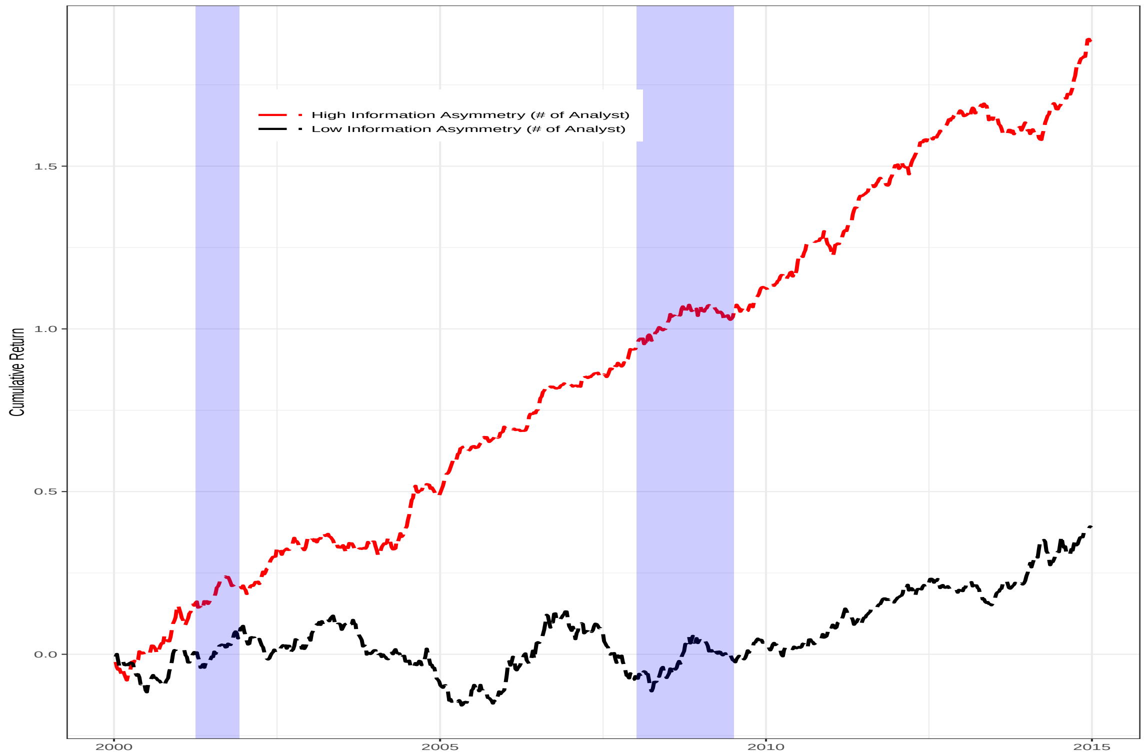


Figure 4: **CIS-based Long-Short Portfolio Performance in Subsamples** This figure shows the cumulative returns of the long-short CIS portfolio in subsamples with different information environments and arbitrage costs. The information environment measures include firm size, analyst coverage, and analyst forecast dispersion, and the arbitrage costs measures are liquidity and return volatility. At the end of each week, we sort all stocks with negative CIS into deciles and form an equal-weighted long-short portfolio by shorting the stocks with the most negative CIS and longing the stocks with the least negative CIS. We then hold the portfolio for 1 week and rebalance at the end of each week. The sample period runs from January 2000 to December 2014.

Figure 4 (continued)



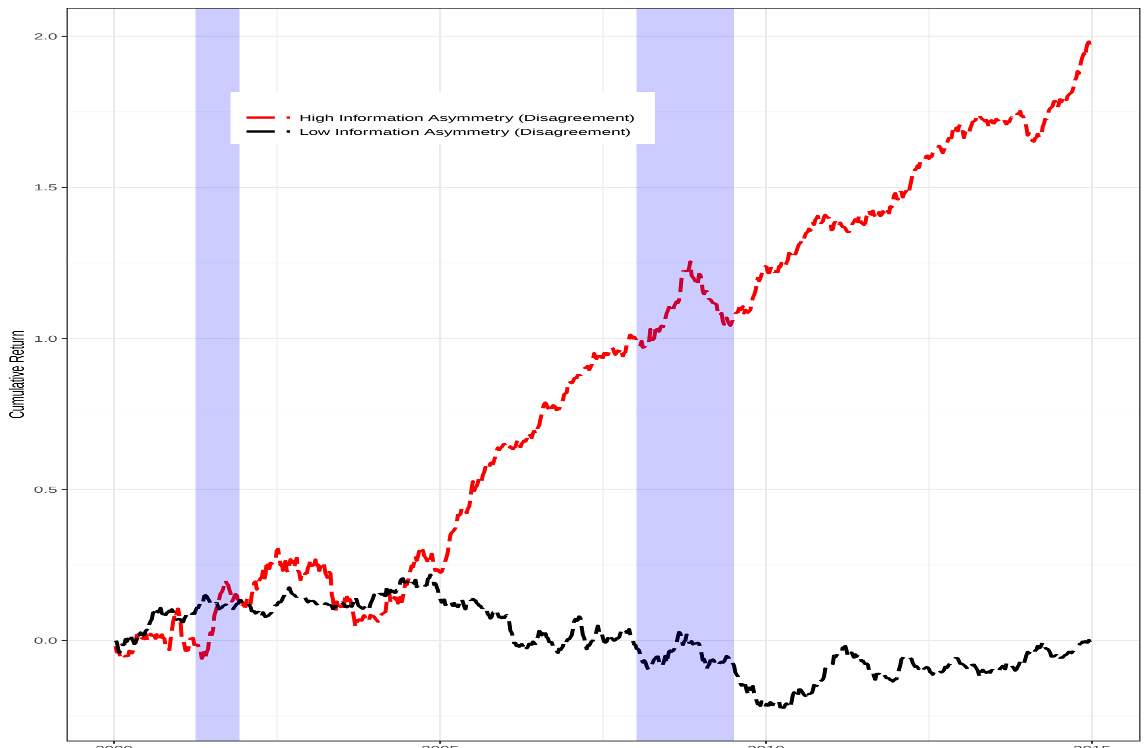
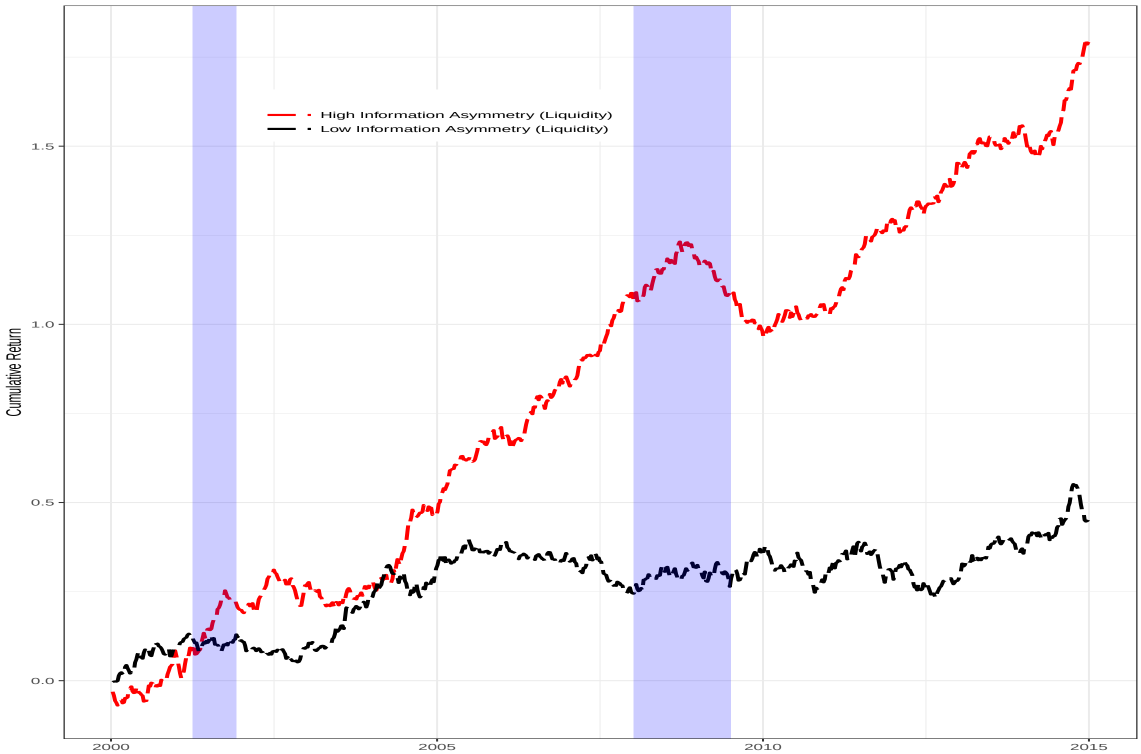


Figure 4 (continued)



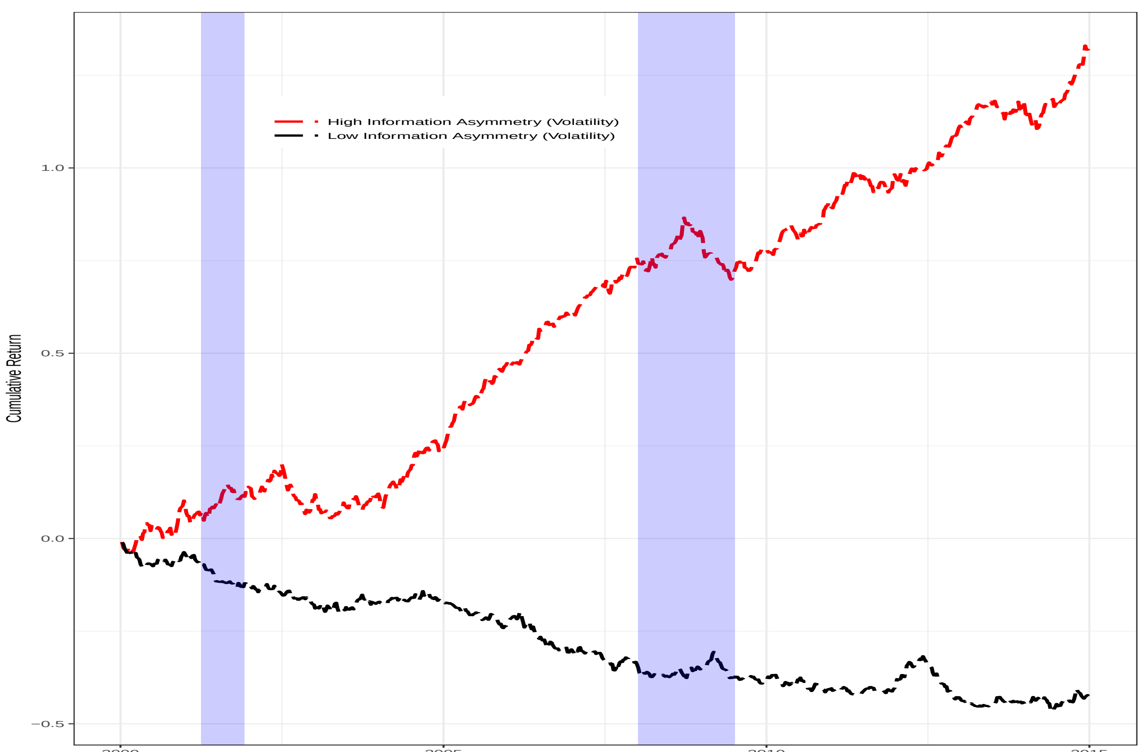


Table 1: **Summary Statistics**

This table reports the summary statistics for the main variables used in the paper. All variables are defined in section 2. Panel A reports the network centrality for Fama-French 30 industries and the price delayness measure of Hou and Moskowitz (2005). Panel B presents weekly industry returns, firm characteristics and Cross-Industry News Signal (CIS). Panel C reports the average cross-industry news tone for Fama-French 30 industries.

Panel A: Cross Industry Media Connection

	EigenCtr	DegreeCtr	Delay	Freq. in Low EigenCtr	Freq. in High EigenCtr	Freq. in Low DegreeCtr	Freq in High DegreeCtr
Util	1.55%	1.97%	11.20%	0	0	0	0
Fin	11.31%	13.13%	0.75%	0	19	0	19
Food	1.78%	2.23%	8.24%	0	0	0	0
Whlsl	3.43%	3.49%	1.56%	0	4	0	5
BusEq	15.05%	14.59%	0.85%	0	19	0	19
Servs	13.07%	10.36%	0.54%	0	19	0	19
Hlth	5.27%	4.86%	1.73%	0	12	0	11
FabPr	2.66%	2.69%	0.88%	0	0	0	1
Games	3.95%	3.40%	3.52%	0	4	0	1
Rtail	6.63%	6.93%	1.17%	0	15	0	15
Cnstr	2.63%	2.44%	0.95%	0	1	0	0
ElcEq	2.20%	1.81%	0.59%	0	0	0	0
Carry	2.54%	2.70%	2.46%	0	0	0	0
Hshld	4.22%	4.47%	4.50%	0	5	0	8
Smoke	0.67%	0.91%	14.24%	12	0	12	0
Chems	1.25%	1.70%	3.15%	0	0	0	0
Books	1.71%	1.48%	2.34%	1	1	1	0
Steel	1.73%	2.00%	1.76%	0	0	0	0
Trans	2.03%	2.22%	1.26%	0	0	0	0
Telcm	5.60%	4.45%	1.23%	0	7	0	7
Oil	3.67%	4.66%	7.12%	0	6	0	8
Clths	0.61%	0.64%	1.77%	11	0	12	0
Txtls	0.08%	0.10%	3.74%	19	0	19	0
Meals	1.37%	1.32%	2.85%	2	0	1	0
Autos	3.25%	3.23%	2.22%	0	2	0	1
Mines	0.57%	0.68%	20.58%	11	0	12	0
Other	0.06%	0.07%	1.74%	19	0	19	0
Beer	0.14%	0.20%	9.70%	19	0	19	0
Coal	0.13%	0.17%	13.54%	18	0	18	0
Paper	0.86%	1.08%	1.85%	2	0		

Panel B: Cross Sectional Return Predictability

	Mean	SD	5%	25%	Median	75%	95%
Return (%)	-0.51	7.26	-12.41	-3.23	-0.00	2.44	10.40
CIS (%)	-0.43	0.74	-1.56	-0.49	-0.21	-0.08	-0.01
Peer News	0.041	0.017	0.00	0.037	0.043	0.051	0.63
Firm News	0.012	0.023	0.00	0.00	0.001	0.012	0.68
# of Cross News	7,436.69	5,180.39	0.00	3,468	7,977	10,890	15,698
# of Peer News	814.44	1,051.30	0.00	138	473	1,063	3,000
# of Firm News	2.42	11.02	0.00	0.00	0.00	1.00	12.00
Size	3.56	3.34	0.00	0.00	3.84	6.18	9.25
B/M	1.05	33.64	0.00	0.00	0.90	1.27	2.95
Turnover	7.38	6.04	0.00	0.00	11.04	12.58	13.80
Leverage	0.14	0.23	0.00	0.00	0.00	0.20	0.67
Volatility*100	0.75	8.54	0.00	0.01	0.06	0.34	2.56

Table 1 (continued)

Panel C: Earnings Annoucement	Mean	SD	5%	25%	Median	75%	95%
Cross Industry News Tone							
Food	0.044	0.006	0.037	0.041	0.043	0.048	0.056
Beer	0.041	0.007	0.032	0.036	0.040	0.045	0.055
Smoke	0.052	0.006	0.044	0.048	0.051	0.056	0.064
Games	0.044	0.006	0.037	0.040	0.043	0.048	0.056
Books	0.040	0.009	0.029	0.035	0.039	0.044	0.057
Hshld	0.043	0.006	0.035	0.039	0.041	0.046	0.056
Clths	0.038	0.008	0.028	0.033	0.038	0.043	0.052
Hlth	0.049	0.005	0.041	0.046	0.049	0.052	0.057
Chems	0.044	0.007	0.035	0.039	0.042	0.047	0.058
Txtls	0.042	0.011	0.028	0.034	0.041	0.049	0.061
Cnstr	0.043	0.007	0.034	0.037	0.042	0.049	0.056
Steel	0.046	0.006	0.038	0.042	0.045	0.049	0.059
FabPr	0.042	0.008	0.032	0.036	0.040	0.048	0.058
ElcEq	0.042	0.008	0.031	0.037	0.040	0.046	0.058
Autos	0.049	0.008	0.038	0.042	0.047	0.053	0.065
Carry	0.041	0.006	0.031	0.036	0.040	0.044	0.051
Mines	0.049	0.008	0.036	0.042	0.049	0.055	0.062
Coal	0.038	0.011	0.023	0.031	0.038	0.044	0.058
Oil	0.050	0.006	0.041	0.046	0.050	0.053	0.059
Util	0.040	0.006	0.032	0.037	0.039	0.043	0.054
Telcm	0.042	0.006	0.034	0.037	0.041	0.046	0.055
Servs	0.043	0.006	0.036	0.039	0.041	0.046	0.056
BusEq	0.043	0.008	0.035	0.038	0.040	0.046	0.061
Paper	0.043	0.008	0.033	0.038	0.041	0.047	0.058
Trans	0.045	0.007	0.036	0.039	0.043	0.049	0.058
Whlsl	0.041	0.006	0.034	0.037	0.040	0.045	0.052
Rtail	0.047	0.006	0.038	0.043	0.045	0.050	0.058
Meals	0.043	0.007	0.033	0.038	0.042	0.048	0.056
Fin	0.046	0.006	0.037	0.041	0.044	0.050	0.057
SUE	0.22	1.48	-1.84	-0.53	0.12	0.85	2.60
Other Variables							
Firm Tone	0.040	0.021	0.008	0.024	0.039	0.054	0.077
Industsry Ttone	0.045	0.009	0.032	0.039	0.044	0.050	0.060
# of Firm News	28.46	57.53	1.00	4.00	11.00	29.00	105.00
# of Industry News	1,237.61	1,582.90	70	277	636	1,446	5,270
# of Cross Industry News	57,506.56	26,226.19	17,900	30,621	64,594	78,568	96,313
Forecast Dispersion	0.04	0.06	0.00	0.01	0.02	0.04	0.13
Forecast Revision	-0.00	0.00	-0.01	-0.00	0.00	0.00	0.00
Size	7.70	2.49	0.00	6.78	7.99	9.22	11.00
B/M	1.75	1.25	0.00	1.08	1.40	2.05	4.01
Turnover	13.56	3.41	0.00	13.81	14.34	14.83	15.54
$AR_{t-252,t-31}$	-0.03	0.17	-0.33	-0.11	-0.02	0.06	0.20
$AR_{t-30,t-3}$	-0.27	10.26	-15.72	-4.36	0.17	4.48	14.41
AR_{t-2}	0.05	2.09	-3.10	-0.88	0.02	0.96	3.28
Consensus Forecast	0.46	0.67	-0.10	0.16	0.35	0.63	1.37
Management Forecast	0.25	0.43	0.00	0.00	0.00	0.00	1.00
Volatility	0.11	0.05	0.06	0.08	0.10	0.14	0.21
Market Return	0.01	0.05	-0.09	-0.02	0.01	0.04	0.08
Institutional Ownership	0.69	0.20	0.30	0.57	0.71	0.83	0.97
Leverage	0.20	0.19	0.00	0.03	0.14	0.31	0.58
Momentum	-0.00	0.30	-0.54	-0.12	0.03	0.16	0.41

Table 2: **Media-Based Industry Centrality and Price Delayness**

In this table, we sort all Fama-French 30 industries into five quintiles based on network centrality measure and report the average price delayness measure of Hou and Moskowitz (2005) for each group from 1996 to 2014. Column 1 (2) shows the eigenvector (degree) centrality, constructed based on media news. Statistical significance of the difference between the highest and lowest centrality quintiles is reported by Newey-West adjusted t-statistics.

	Eigen-centrality	Degree-centrality
Low Centrality of Industry in Media Network	7.90%	7.62%
2	5.44%	4.01%
3	2.24%	4.51%
4	2.60%	2.90%
High Centrality of Industry in Media Network	1.32%	1.40%
High - Low	-6.58%	-6.22%
T-stats	-3.96	-4.17

Table 3: **Pairwise Industry Connection and Cross-Industry Information Delayness**

This table reports the cross-industry information delayness for five groups sorted on pairwise industry connection. Our measure of pairwise industry connection is the number of news mentioning an industry pair simultaneously. Similar to the delay measure of Hou and Moskowitz (2005), for an industry pair A and B, the delayness of industry B's news on industry A's return, $Delay_{B \to A}$, is the fraction of industry A's returns explained by industry B's lagged returns. More specifically, the measure is one minus the ratio of the R^2 from regression (3.1) by restricting $\delta_j^{-n} = 0, n \in [1,\ 4]$, over the R^2 from regression (3.1) without restrictions. The pairwise information delayness between industry A and B is the average of $Delay_{A \to B}$ and $Delay_{B \to A}$. We then sort all industry pairs into five quintiles according to the # of news mentioning the industry pair simultaneously and report the average delayness of each quintile. We test the statistical significance of the difference between the highest and lowest quintiles of Pairwise Industry Connection and report the Newey-West adjusted t-statistics.

	$Average\ Delay\ _{A \leftrightarrow B}$	$Delay_{A \to B}$	$Delay_{B \to A}$
Low # of Connected News between A and B	13.21%	13.40%	13.01%
2	9.97%	10.30%	9.65%
3	6.78%	6.70%	6.85%
4	6.26%	5.89%	6.64%
High # of Connected News between A and B	5.06%	4.85%	5.27%
High - Low	-8.15%	-8.56%	-7.75%
T-stats	-5.14	-5.00	-5.17

Table 4: **Cross-Industry News and Earnings Surprise**

This table reports the predictability of cross-industry news for earnings surprise (SUE). The regression is run as follows:

$$\text{SUE}_{it} = \alpha_i + \sum_{\mathbf{J}=1}^{N-1} \beta_{\mathbf{J}} \text{Cross Industry News}_{i,\mathbf{J},t-90,t-3} + \gamma^{'}\text{X} + \epsilon_{it},$$

The dependent variable, SUE, is the standardized unexpected earnings following Bernard and Thomas (1989). Cross Industry News$_{i,\mathbf{j},t-90,t-3}$ is the news tone of Industry J measured over the period (t-90, t-3) relative to the earnings announcement day t. X denotes other explanatory variables. We only include the news tone of one cross industry in Panel A, and all cross-industry news in Panel B for the regression. For each panel, we have 3 specifications with different control variables. The first three columns show the estimated coefficients, T-value and adjusted R^2 for the univariate regression. The middle three columns report the corresponding results that follow the specification of Tetlock *et al.* (2008). In the last three columns, we added all control variables. Standard errors are clustered at firm level.

Panel A: One Industry	SUE								
Empirical Design:	Univariate			Tetlock 2008			All Controls		
	Coef	T-value	R^2 (%)	Coef	T-value	R^2 (%)	Coef	T-value	R^2 (%)
Food	-0.08	-9.68	0.30	-0.05	-5.66	17.41	-0.02	-1.77	16.72
Beer	-0.10	-12.76	0.52	-0.06	-7.46	17.55	-0.04	-3.93	16.89
Smoke	-0.04	-4.60	0.07	-0.02	-2.24	17.31	-0.00	-0.16	16.72
Games	-0.14	-16.54	0.88	-0.08	-10.36	17.60	-0.05	-4.97	16.79
Books	-0.13	-16.31	0.84	-0.08	-10.71	17.57	-0.06	-6.49	16.85
Hshld	-0.18	-21.80	1.52	-0.11	-13.56	17.78	-0.09	-8.12	16.96
Clths	-0.10	-12.25	0.48	-0.04	-5.93	17.23	-0.02	-2.47	16.57
Hlth	-0.05	-5.73	0.11	-0.03	-4.03	18.13	0.00	0.30	17.55
Chems	-0.12	-14.70	0.71	-0.07	-8.49	17.31	-0.04	-3.89	16.61
Txtls	-0.12	-14.96	0.71	-0.07	-8.85	17.56	-0.05	-4.56	16.83
Cnstr	-0.15	-18.23	1.09	-0.09	-11.52	17.16	-0.06	-5.16	16.34
Steel	-0.12	-14.72	0.70	-0.06	-8.52	17.37	-0.03	-3.42	16.66
FabPr	-0.15	-17.94	1.06	-0.08	-10.26	17.29	-0.06	-5.45	16.46
ElcEq	-0.15	-17.86	1.02	-0.09	-11.04	17.67	-0.08	-7.12	16.91
Autos	-0.18	-21.69	1.50	-0.12	-14.87	17.74	-0.11	-10.40	16.98
Carry	-0.12	-14.30	0.65	-0.06	-7.37	17.36	-0.03	-3.11	16.64
Mines	-0.08	-10.10	0.33	-0.05	-6.65	17.55	-0.03	-3.78	16.90
Coal	0.06	7.17	0.19	0.03	4.15	15.68	0.04	3.87	15.89
Oil	-0.05	-6.05	0.12	-0.03	-4.05	17.74	0.01	0.68	17.20
Util	-0.04	-5.03	0.09	-0.02	-2.14	18.63	0.03	2.68	18.14
Telcm	-0.12	-14.11	0.64	-0.07	-8.69	17.77	-0.03	-2.54	16.99
Servs	-0.12	-14.68	0.73	-0.07	-9.00	17.11	-0.04	-3.70	16.37
BusEq	-0.13	-14.25	0.71	-0.07	-8.77	17.01	-0.06	-4.95	16.22
Paper	-0.13	-15.99	0.82	-0.07	-9.22	17.61	-0.04	-3.37	16.82
Trans	-0.18	-22.09	1.60	-0.11	-13.92	17.47	-0.09	-8.45	16.55
Whlsl	-0.13	-15.90	0.82	-0.07	-8.95	17.22	-0.05	-4.41	16.51
Rtail	-0.15	-17.74	1.07	-0.09	-11.04	17.00	-0.07	-6.65	16.14
Meals	-0.11	-12.90	0.54	-0.06	-7.58	17.35	-0.04	-3.36	16.69
Fin	-0.08	-9.33	0.34	-0.04	-4.86	17.51	0.00	0.22	17.04
Other	-0.09	-10.47	0.35	-0.05	-6.48	17.45	-0.04	-3.89	16.82
Year effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4 (continued)

Panel B: All Industries	SUE					
Empirical Design:	No Other Controls		Tetlock 2008		All Controls	
	Coef	T-value	Coef	T-value	Coef	T-value
Food	0.06	2.44	0.04	1.79	0.06	2.23
Beer	-0.05	-4.83	-0.03	-2.61	-0.03	-2.53
Smoke	-0.06	-3.15	-0.06	-3.51	-0.06	-2.66
Games	-0.08	-3.56	-0.04	-1.95	-0.05	-1.86
Books	-0.03	-1.94	-0.03	-2.05	-0.03	-2.05
Hshld	-0.03	-1.29	-0.04	-1.61	-0.03	-1.22
Clths	0.03	2.10	0.03	1.95	0.02	1.08
Hlth	0.18	6.50	0.10	3.93	0.11	3.70
Chems	-0.01	-0.30	-0.00	-0.15	-0.00	-0.21
Txtls	-0.04	-2.94	-0.01	-0.95	-0.01	-0.64
Cnstr	-0.04	-2.30	-0.03	-1.98	-0.01	-0.48
Steel	-0.04	-2.01	-0.01	-0.35	0.00	0.02
FabPr	0.01	0.28	0.02	1.26	0.03	1.68
ElcEq	-0.08	-4.25	-0.06	-3.53	-0.07	-3.56
Autos	-0.16	-8.44	-0.12	-7.35	-0.13	-6.83
Carry	0.04	1.79	0.04	2.15	0.04	1.79
Mines	-0.04	-2.75	-0.03	-2.16	-0.03	-2.25
Coal	0.05	4.59	0.01	0.93	0.01	0.93
Oil	0.09	4.44	0.02	1.32	0.01	0.71
Util	0.10	4.35	0.06	2.72	0.05	1.81
Telcm	-0.01	-0.34	0.01	0.48	0.03	1.14
Servs	0.07	2.94	0.04	1.80	0.04	1.39
BusEq	0.06	2.64	0.05	2.24	0.04	1.75
Paper	0.03	1.59	0.02	1.37	0.06	2.48
Trans	-0.13	-5.35	-0.09	-3.82	-0.08	-3.12
Whlsl	0.00	0.07	0.01	0.30	-0.02	-0.63
Rtail	-0.00	-0.11	0.00	0.15	0.02	0.71
Meals	0.06	3.40	0.04	2.51	0.03	1.44
Fin	0.05	1.74	0.03	1.05	0.04	1.23
Other	-0.02	-1.69	-0.02	-1.51	-0.01	-1.23
Year effect	Yes		Yes		Yes	
Firm effect	Yes		Yes		Yes	
N	32,917		32,917		28,206	
adj. R^2 (%)	2.59		18.25		17.47	

Table 5: **Fama-MacBeth regressions of stock returns on CIS**

This table reports the Fama-MacBeth regression of stock returns on cross-industry news signals (CIS). CIS is the out-of-sample forecasted return based on cross-industry-news tones. Peer News is average news tone of peer firms within the same industry. Firm News is the firm-specific news tone. We only include stocks with negative CIS in the regression. The sample period is from Jan 2000 to Dec 2014. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	2000 - 2014			2000 - 2008			2009 - 2014		
CIS	0.137***	0.130***	0.095***	0.126***	0.118***	0.079***	0.153***	0.148***	0.118***
	(7.15)	(6.48)	(5.09)	(5.21)	(4.51)	(3.32)	(4.90)	(4.74)	(3.98)
Lagged Return	-0.036***	-0.036***	-0.042***	-0.039***	-0.039***	-0.045***	-0.032***	-0.032***	-0.038***
	(-21.08)	(-22.12)	(-23.60)	(-17.62)	(-18.46)	(-19.27)	(-11.81)	(-12.43)	(-13.73)
Peer News		-0.001	-0.002		-0.001	-0.001		-0.001	-0.002
		(-0.60)	(-1.53)		(-0.56)	(-1.14)		(-0.28)	(-1.02)
Firm News		-0.001**	-0.002***		-0.000	-0.001***		-0.001***	-0.002***
		(-2.04)	(-5.02)		(-0.68)	(-2.72)		(-2.63)	(-5.09)
# of Peer News		0.000	0.000		0.000	-0.000		-0.000	0.000
		(0.01)	(0.24)		(0.06)	(-0.54)		(-0.07)	(1.11)
# of Firm News		0.001***	0.000		0.000*	-0.000**		0.001***	0.001***
		(5.17)	(0.76)		(1.66)	(-2.25)		(7.29)	(6.05)
Size			0.001***			0.001***			0.001***
			(8.39)			(6.67)			(5.13)
B/M			0.000***			0.000***			0.000***
			(8.19)			(6.10)			(5.52)
Turnover			-0.000***			-0.000***			-0.000***
			(-4.83)			(-3.84)			(-3.00)
Leverage			-0.003***			-0.003***			-0.002***
			(-5.90)			(-4.58)			(-3.78)
Volatility			-0.013***			-0.009*			-0.020***
			(-3.41)			(-1.73)			(-3.44)
Intercept	-0.000	-0.000	-0.001**	-0.001	-0.001	-0.001***	0.001	0.001	-0.000
	(-0.04)	(-0.14)	(-2.41)	(-0.94)	(-1.25)	(-2.95)	(0.99)	(0.82)	(-0.64)
N	1,401,162	1,401,162	1,401,162	855,092	855,092	855,092	546,070	546,070	546,070
Average R^2 (%)	1.17	2.43	4.00	1.20	2.54	4.06	1.14	2.28	3.91

Table 6: **Portfolio Alphas Sorted by Cross-Industry News Signal**

This table reports the weekly alpha of the long-short portfolio constructed on cross-industry news signal (CIS). At the end of each week, we sort all stocks with negative CIS into deciles and form an equal-weight long-short portfolio by shorting the stocks with most negative CIS and longing the stocks with the least negative CIS. We then hold the portfolio for one week and rebalance the portfolio at the end of each week. Column (1), (4) and (7) report the CAPM adjusted alpha, column (2), (5) and (8) for the Fama-French three factor adjusted alpha, and column (3), (6) and (9) for the Carhart (1997) four-factor adjusted alpha. Standard errors are computed using the White (1980) heteroskedasticity-consistent covariance matrix. ***, ** and * indicate significance at 1%, 5%, and 10% levels, respectively.

	2000 - 2014			2000 - 2008			2009 - 2014		
Alpha (%)	0.21***	0.22***	0.21***	0.21***	0.22***	0.21***	0.21***	0.21***	0.22***
	(5.17)	(5.31)	(5.22)	(3.44)	(3.61)	(3.37)	(4.47)	(4.53)	(4.75)
Market Risk	-0.04**	-0.03	-0.03*	-0.05**	-0.03	-0.04*	-0.02	-0.01	-0.02
	(-2.49)	(-1.65)	(-1.71)	(-2.21)	(-1.53)	(-1.83)	(-0.99)	(-0.61)	(-0.74)
SMB		-0.11***	-0.13***		-0.12***	-0.15***		-0.08*	-0.09**
		(-3.49)	(-4.12)		(-2.91)	(-3.56)		(-1.83)	(-2.06)
HML		0.03	0.05*		0.01	0.03		0.03	0.07*
		(0.99)	(1.94)		(0.35)	(0.66)		(0.79)	(1.76)
UMD			0.08***			0.09***			0.06***
			(4.76)			(3.99)			(2.70)
N	772	772	772	462	462	462	310	310	310
adj. R^2	0.007	0.025	0.052	0.008	0.026	0.057	-0.000	0.011	0.031

Table 7: **Transaction Cost and the Profitability of CIS-based Strategy**

This table shows the impact of transaction costs on the profitability of long-short strategy based on cross-industry news and firm-specific news. Reported is the (annualized) raw returns and Fama-French three-factor alphas under the assumption that a trader must incur a round-trip transaction cost varing from 1 to 10 bps.

Trading	Cross Industry News			Firm Specific News		
Cost (bps)	Raw Return (%)	α(%)	T_α	Raw Return (%)	α (%)	T_α
1	8.89	9.20	4.12	7.89	7.86	3.57
2	8.30	8.60	3.86	7.30	7.27	3.30
3	7.71	8.01	3.59	6.71	6.68	3.03
4	7.11	7.42	3.32	6.12	6.09	2.76
5	6.53	6.83	3.07	5.52	5.50	2.49
6	5.94	6.24	2.80	4.93	4.91	2.22
7	5.35	5.65	2.54	4.34	4.32	1.95
8	4.76	5.06	2.27	3.75	3.73	1.68
9	4.17	4.47	2.01	3.16	3.14	1.41
10	3.58	3.88	1.74	2.57	2.55	1.14

Table 8: **Persistence of CIS-based Strategy**

This table shows the persistence of the CIS-based strategy. At the end of each week, we form a long-short portfolio based on cross-industry news or firm-specific news observed 2 to 10 weeks ago, and hold the portfolio for 1 week, and rebalance weekly. Reported is the (annualized) raw returns and Fama-French three-factor alphas of the long-short portfolio.

Week	Cross Industry News				Firm Specific News			
after News	Raw Return (%)	T_{Raw}	α (%)	T_α	Raw Return (%)	T_{Raw}	α (%)	T_α
2	11.49	5.65	13.16	6.63	2.96	1.70	3.12	1.85
3	9.40	4.59	11.28	5.60	2.99	1.63	3.19	1.94
4	10.77	5.31	12.25	6.18	3.82	2.18	3.87	2.30
5	13.01	5.81	14.84	7.11	4.58	2.67	5.13	3.11
6	10.14	4.97	12.43	6.28	1.91	1.09	2.25	1.32
7	10.48	4.96	12.69	6.08	1.43	0.84	1.86	1.14
8	11.97	5.46	13.65	6.33	1.18	1.49	1.38	1.68
9	13.77	6.63	15.79	7.79	3.82	2.28	3.97	2.44
10	9.86	4.48	10.85	4.99	1.69	1.05	2.05	1.31

Table 9: **Performance of CIS-based Strategy after Controlling for Alternative Information**

This table reports the weekly alpha of the long-short portfolio constructed on cross-industry news signal (CIS). At the end of each week, we sort all stocks with negative CIS into deciles and form an equal-weight long-short portfolio by shorting the stocks with most negative CIS and longing the stocks with the least negative CIS. We then hold the portfolio for one week and rebalance the portfolio at the end of each week. Column (1), (4) and (7) reports the alphas of CIS-based strategy after adding the portfolio returns based on the news of within-industry peer. Column (2), (5) and (8) reports the alphas after adding portfolio returns based on firms' own news. Columns (3), (6) and (9) reports the alphas after adding portfolio returns based on cross-industry returns. ***, ** and * indicate significance at 1%, 5%, and 10% levels, respectively.

	2000 - 2014			2000 - 2008			2009 - 2014		
Alpha (%)	0.20***	0.15***	0.14***	0.20***	0.16***	0.15**	0.20***	0.14***	0.13***
	(4.95)	(3.79)	(3.45)	(3.23)	(2.64)	(2.46)	(4.42)	(3.00)	(2.79)
Market Risk	-0.02	-0.03*	-0.03*	-0.04	-0.04*	-0.04	-0.02	-0.02	-0.03
	(-1.55)	(-1.79)	(-1.88)	(-1.63)	(-1.69)	(-1.62)	(-0.71)	(-0.80)	(-1.42)
SMB	-0.12***	-0.11***	-0.13***	-0.14***	-0.12***	-0.16***	-0.09**	-0.10**	-0.09**
	(-4.05)	(-3.76)	(-4.40)	(-3.42)	(-3.06)	(-3.89)	(-2.18)	(-2.36)	(-2.10)
HML	0.04	0.06**	0.04	0.02	0.04	0.01	0.06	0.07*	0.07*
	(1.61)	(2.20)	(1.60)	(0.53)	(0.99)	(0.30)	(1.44)	(1.77)	(1.87)
UMD	0.06***	0.02	0.01	0.07***	0.03	0.03	0.04*	-0.01	-0.01
	(3.60)	(0.88)	(0.81)	(3.14)	(1.13)	(1.39)	(1.81)	(-0.23)	(-0.63)
Peer News	0.10***	0.07***	0.06**	0.09**	0.07*	0.05	0.10***	0.06*	0.06*
	(3.51)	(2.61)	(2.15)	(2.26)	(1.86)	(1.39)	(3.09)	(1.92)	(1.74)
Firm News		0.49***	0.44***		0.47***	0.40***		0.52***	0.50***
		(5.89)	(5.37)		(3.89)	(3.38)		(5.30)	(5.04)
Cross-Industry Return			0.26***			0.30***			0.19***
			(5.92)			(5.03)			(2.94)
N	772	772	772	462	462	462	310	310	310
adj. R^2	0.065	0.105	0.143	0.065	0.093	0.139	0.057	0.134	0.156

Table 10: **Economic Uncertainty and Profitability of CIS-based Strategy**

This table reports the raw returns and risk-adjusted returns of CIS-based portfolio over periods of high and low economic uncertainty. Our proxies for economic uncertainty include VIX, Economic Policy Uncertainty (EPU), and a measure of market-wide news dispersion. News dispersion is defined as the cross-sectional standard deviation of news tone across firms. A period is indicated as high (low) uncertainty if the economic uncertainty index in the previous week is above (below) the median value of the whole sample. The sample period is between Jan 2000 and Dec 2014. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Average	Annulized Risk Adjusted Return			
	Return	CAPM	FF3	FF3M	All Control
Panel A: VIX					
Low	0.07**	0.08**	0.08**	0.09**	0.06*
	2.54	2.26	2.38	2.50	1.87
High	0.15***	0.15***	0.15***	0.15***	0.11***
	4.53	3.52	3.65	4.07	3.23
Panel B: EPU					
Low	0.11***	0.12***	0.12***	0.12***	0.08***
	3.52	3.27	3.31	3.29	2.69
High	0.11***	0.11***	0.11***	0.12***	0.09***
	3.66	2.95	3.21	3.48	2.73
Panel C: News Dispersion					
Low	0.06**	0.07*	0.08**	0.08**	0.06*
	2.11	1.88	2.12	2.18	1.72
High	0.16***	0.16***	0.16***	0.15***	0.12***
	4.87	4.06	4.09	4.40	3.36

Table 11: **Cross-industry News and Analyst Forecast**

This table presents results from panel regression of forecast revision or improvement in forecast accuracy on Cross-industry News Tone. The regression is run as follows:

$$\text{Y}_{ijt} \quad = \quad \alpha + \beta_1 \text{Average Cross News Tone}_{t-90,t-3} + \gamma^{`}\text{X} + \epsilon_{ijt},$$

where Y_{ijt} is the forecast revision or improvement in forecast accuracy. Forecast revision is defined as the absolute change of two adjacent forecasts scaled by stock price at the end of the previous year. Forecast accuracy improvement is defined as the change in accuracy in two adjacent forecasts, where forecast accuracy is defined as the negative absolute value of the difference between actual earnings and forecasted earnings. X includes a set of explanatory variables, defined previously. Standard errors are clustered at the firm level. ***, ** and * indicate significance at 1%, 5%, and 10% levels, respectively.

Table 11 (continued)

Dependent Variable:	**Forecast Revision:** $\frac{\lvert\Delta\text{Forecast Value}\rvert}{Price}$			**Forecast Improvement:** $\frac{\Delta Accuracy}{Price}$		
Cross News Tone	0.01984***	0.00989***	0.00962***	0.03283***	0.01786**	0.01862**
	(7.09)	(3.00)	(3.26)	(3.71)	(2.08)	(2.09)
# of Cross Industry News	0.00000	0.00000	0.00003**	0.00001	0.00003	0.00006*
	(0.02)	(0.09)	(2.21)	(0.44)	(0.94)	(1.67)
Firm News Tone		0.13789***	0.08705***		0.14815***	0.07621***
		(8.19)	(7.64)		(4.15)	(2.71)
# of Firm News		0.00083***	0.00056***		0.00083***	0.00027*
		(5.39)	(3.73)		(4.18)	(1.66)
Industry News Tone		0.00797***	0.00547***		0.01307***	0.00849***
		(4.22)	(3.26)		(3.70)	(2.69)
# of Industry News		-0.00049***	-0.00040***		-0.00073***	-0.00058***
		(-4.43)	(-4.19)		(-3.23)	(-2.95)
Analyst Dispersion			0.01400***			0.01852***
			(9.05)			(2.91)
Forecast Revision			-0.15738***			-0.42410**
			(-5.38)			(-2.53)
Size			-0.00474**			-0.00447
			(-2.07)			(-0.93)
B/M			0.01290			0.01551
			(1.05)			(0.71)
Turnover			-0.00008			-0.00033
			(-0.26)			(-0.49)
$AR_{t-252,t-31}$			-0.01633**			-0.01799
			(-2.18)			(-0.67)
$AR_{t-30,t-3}$			0.00002			-0.00004*
			(1.48)			(-1.71)
AR_{t-2}			-0.00007			-0.00021**
			(-1.19)			(-2.13)
Consensus Forecast			-0.00073			0.00771**
			(-1.45)			(1.97)
Analyst Boldness			0.00258***			0.00128***
			(19.60)			(4.51)
Forecast Horizon			0.00000***			0.00001***
			(5.18)			(3.90)
Forecast Frequency			0.00561**			0.00231
			(1.97)			(0.73)
Genearl Exp			-0.00000			0.00000
			(-0.10)			(0.91)
Firm Exp			0.00000*			-0.00000
			(1.81)			(-1.40)
Firm Coverage			-0.00001***			-0.00001
			(-2.95)			(-0.59)
Analyst Ranking			0.00001***			0.00009***
			(2.82)			(8.83)
Abnormal # of Analysts			-0.00000			-0.00000
			(-1.64)			(-0.96)
Earnings Surprise			-0.00068***			0.00087
			(-3.01)			(0.80)
Return Volatility			0.12953*			0.01076
			(1.71)			(0.08)
Market Return			-0.02520**			-0.04724*
			(-2.34)			(-1.87)
Institutional Ownership			-0.01141			-0.02579
			(-1.17)			(-1.53)
Leverage			0.00888***			0.01048***
			(7.42)			(3.22)
Momentum			-0.00872**			-0.04863*
			(-2.26)			(-1.90)
Illiquidity			805.17221***			-316.44915
			(3.62)			(-0.47)
Overconfidence			0.00006			0.00201***
			(0.39)			(3.97)
Intercept	-0.00347***	-0.00289***	-0.03959***	-0.00843***	-0.00644**	-0.08320**
	(-3.88)	(-2.91)	(-3.13)	(-2.96)	(-2.24)	(-2.33)
Year effect	Yes	Yes	Yes	Yes	Yes	Yes
Firm effect	Yes	Yes	Yes	Yes	Yes	Yes
N	93,946	93,109	93,090	93,913	93,076	93,057
adj. R^2	0.322	0.330	0.435	0.057	0.058	0.090

Table 12: **Fama-MacBeth Regressions of Firm-Specific News on CIS**

This table reports the Fama-MacBeth regression of firm-specific news tone on lagged Cross-industry news signals(CIS), controlling for lagged stock returns and other firm characteristics. The sample runs from Jan 2000 to Dec 2014. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:	Firm-specific News Tone		
	2000 - 2014	2000 - 2007	2008 - 2014
CIS	-0.42***	-0.53***	-0.30***
	(-15.81)	(-13.22)	(-9.11)
Peer News	0.11***	0.12***	0.09***
	(32.83)	(29.24)	(18.51)
Lagged Dependent Variable	0.14***	0.15***	0.13***
	(64.15)	(77.21)	(32.67)
# of Peer News	-0.00***	-0.00***	-0.00***
	(-25.08)	(-20.40)	(-20.27)
# of Firm News	0.01***	0.01***	0.00***
	(40.69)	(36.80)	(27.64)
Lagged Return	-0.09***	-0.09***	-0.09***
	(-24.39)	(-18.76)	(-15.92)
Size	0.01***	0.02***	0.01***
	(47.93)	(41.79)	(28.41)
B/M	-0.00***	-0.00***	-0.00***
	(-7.26)	(-3.46)	(-6.76)
Turnover	0.00***	0.00***	0.00***
	(28.00)	(21.13)	(18.88)
Leverage	-0.02***	-0.02***	-0.02***
	(-19.88)	(-16.05)	(-12.08)
Volatility	0.05***	0.02**	0.07***
	(6.48)	(2.52)	(6.51)
Intercept	-0.05***	-0.05***	-0.05***
	(-24.98)	(-22.51)	(-14.24)
N	458,479	281,026	177,453
Average R^2 (%)	25.27	25.38	25.12