Media News and Cross Industry Information Diffusion

Li GUO

Singapore Management University

June 2017

Abstract

Media news serves as information intermediary that contributes to the cross industry return predictability. First, we find that cross industry news contains valuable information about firm fundamentals that is not priced by the market. Second, consistent with high information costs hypothesis, cross industry news has long term effects on future returns with an annulized risk adjusted return 10.85% after 10 weeks of the signal. Third, cross industry news is more valuable to small stocks, illiquid stocks, and stocks with high return volatility or low analyst coverage. Fourth, analyst forecasts, institutional fund flows and media news might be the channels that interpret cross industry news to the market. Overall, we provide direct evidence to support the argument that news travels slowly across different industries.

Keywords: Return predictability; information diffusioin; media news; sentiment
JEL Classification: G11; G12; G14

1 Introduction

Information diffusion is widely used to explain cross asset return predictability (Cohen and Frazzini (2008); Menzly and Ozbas (2010); Hong and Stein (1999); Hong and Valkanov (2007), and Rapach (2015)). According to Hong and Valkanov (2007), information diffusion relies on two key assumptions. One is investors' limited attention that suggests many investors may not pay attention to the information from the asset prices of markets that they do not specialize in. The other one suggests that news travels slowly across markets. Indeed, many literature provide empirical evidence and theoretical models to support the limited attention argument[1] while the more important assumption that news travels slowly across markets is under explored.

Indeed, there are other explanations on the cross asset return predictability. For example, a bunch of literature suggests that more liquid stocks lead less liquid stocks, including Lo and MacKinlay (1990), Brennan and Swaminathan (1993), Badrinath and Noe (1995), Jegadeesh and Titman (1995). They find different proxies of liquidity, such as size and analyst coverage et.al and provide empirical evidence to support this argument. On top of that, Boudoukh and Whitelaw argue that many of these lead-lag relationships are due to the own-autocorrelation of portfolios and a high contemporaneous correlation among portfolios. As a result, filling the gap of empirical evidence and key assumption of information diffusion is important for us to accept information diffusion to explain the cross asset return predictability. In this paper, we use Media News as direct measure of cross industry information and shed light on how news travels across different industries.

The argument that news travels slowly across markets can be decomposed into two hypothesis. The first one is that cross industry news contains valuable information about firm fundamentals. Recent work on media news suggests that media news contains soft

[1]Prominent proxies for limited attention in the cross-section include extreme returns, extreme trading volume, and media coverage, such as Ahern and Sosyura (2014); Barber and Odeani (2008); Fang and Peress (2009); Hou and Xiong (2009); Gervais and Mingelgrin (2001); Loh (2010) and Yuan (2015).

information about firms' fundamental values, and so it has incremental explanatory power on firms' future performance, especially when hard information is incomplete or is biased. For example, Tetlock (2008) finds that negative words predict future earnings and Bushee (2010) show that the media serves as an information intermediary which incrementally contribute to firms information environment. However, current literature does not cover the cross industry news information yet. The challenge of this study is the informativeness of cross industry news. In most previous work, the accuracies of individual stock prediction are higher when only company-related news are used as inputs, compared with when sector-related news are used. This is because it is difficult to investigate the relationship among companies, and therefore news about other companies can be noise for predicting the stock prices of a company. In this paper, we construct Cross Industry News Signal (CIS) to solve this problem.

The second hypothesis is that cross industry news slowly diffuses, namely, it has a long term effect on future stock returns. This assumption is built on the spirit of theory model in Merton (1987) and Hong and Stein (1999). From media news perspective, we interpret this assumption as hard interpretation of cross industry news. Different from limited attention assumption, slow news traveling is the case that investors realize the cross industry information while lacking in ability to interpret this information. In this argument, we emphasize the importance of information intermediary that contributes to the news traveling and cross industry news should be more valuable to the stocks with poor information environment.

To answer the question how news travels across industries, we do the textual analysis on the Thomson Reuters News Archive and construct news tone for Fama-French 30 industries. For each news article, the news tone is defined as the proportion of negative words following Tetlock (2008).

We then first study the informativeness of cross industry news by examining its prediction power on Standard Unexpected Earnings (SUE). If cross industry news is valuable to the market, it must contain incremental information on firm fundamentals. Indeed, our first main result shows that cross industry news tone predicts future earnings even controlling for other firm fundamentals and investors' expectation. Similar to previous studies, we find oil as core industry, negatively affects other industry earnings. Meanwhile, Util, Autos and ElcEq also show strong impacts on other industry earnings due to technology spillover effect (Rapach (2015)). Overall, we find firms have complex industry dependencies that may contribute to cross-industry return predictability.

After that, we link cross industry news to cross industry return predictability. Different from previous studies, we do not directly predict industry returns. Instead, we predict future return at firm level and then construct portfolios based on the Cross Industry News Signal (CIS). There are some advantages of doing this. First, for stocks within the same industry, they may react differently toward cross industry news due to complex business model and cross firm connections. In this case, to predict individual stock return, we fully explore the firms' sensitivity to the cross industry news. Second, due to potential measurement error of news tone (Tetlock (2008)) at industry level, to sort industry return may induce forecast error since only 30 industries will be ranked. While sorting on individual stocks, we have 2,233.7 firms per week on average, hence reducing the forecast error in a way.

Our second main result is that stock price responds to the information embedded in negative CIS with a small and long delay. Figure 1 shows the information of negative CIS lasts more than 1 year for market to digest. Meanwhile, we further test the cross sectional premium of negative CIS. Fama-MacBeth regression shows consistent results across different empirical settings and its information is stronger during the most recent years compared to the early years. As a result, we identify potential profits from using weekly trading strategies based on the negative CIS. The trading strategy survives after accounting for reasonable transaction costs and common risk factors. We further consider the forecast horizon of CIS and show that CIS portfolio keeps annualized risk adjusted return 10.85% after 10 weeks of the signal. On the contrary, firm specific news portfolio cannot survive more than 4 weeks. To interpret these results further, we consider potential overlapped information source of CIS, including peer industry news, firm specific news and lagged cross industry returns. Indeed, CIS remains alpha 13.2 bps per week at 1% level with adjusted R^2 15.6%. This suggests that cross industry news contains soft information that is not priced by the current market, consistent with Tetlock (2008).[2] In addition, by controlling lagged return in a predictive regression, our findings are not due to own-autocorrelation of portfolios (Hong and Valkanov (2007)).

We further conduct several analysis to show the robustness of information diffusion story. First, by sorting CIS within small size stocks, our strategy generates higher cumulative returns than sorting on big stocks, consistent with information story that cross industry news reveals more valuable information to small stocks. This finding also applies to stocks shown strong information asymmetry, such as high volatility stocks, illiquid stocks and those with low analyst coverage or high analyst dispersion. On top of that, CIS is more valuable in a high sentiment period. The annualized return is 16% in high sentiment periods while it is only 5% in low sentiment periods, consistent with the spirit of Stambaugh (2012) . Cross industry news also contributes to a better information environment during a high uncertainty period proxied by VIX and news dispersion, while it is not sensitive to policy uncertainty.

In the last, we further point out the channels of news traveling. Due to investors' limited attention and hard interpretation of cross industry news, without information intermediary, it may not be realized by investors. We then consider three potential channels, including analyst forecasts, institutional fund flows and media news. Indeed, we find a large average cross industry news tone significantly affects analyst forecast revisions and improves their forecast accuracy. Moreover, on average, 7.85 cross industry news

[2] We also conduct additional test by running cross industry return based portfolio on CIS and results show that Cross Return based portfolio loses alpha. This result further suggests our measure of cross industry news is reasonable good to explain the cross industry return predictability.

tone show significant effects on the subsequent institutional fund flows and the adjusted R^2 using only cross industry news tones is on average 4.82 times of using only industry fund flow and industry news tone. Last, CIS significantly predicts firm specific news tone which means firm news incorporates cross industry news in a delayed timer.

We see our main contribution in providing direct evidence to fill the gap between the key assumption of information diffusion and its extensions of cross industry return predictability. On top of that, our study links to three mainstream literatures. First, we contribute to cross industry return predictability. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) find that economic links among certain individual firms and industries lead to cross-firm and cross-industry return predictability. Moreover, Hong and Valkanov (2007) show that returns of leading industry significantly predict market index. Different from above studies, we predict individual firms' return to fully explore the sensitivity of firm performance to the cross industry information. In this way, our study is not restricted to the economic linked industries but also take into account of additional industry connections such as technology spill-over effect or firms' business network.

Second, we contribute to the recent and growing literature on how soft information in news can be quantified and linked to asset prices. Tetlock (2007) analyzes the content of a commentary section in the Wall Street Journal, and finds that pessimistic words predict low stock returns. Davis (2006), Engelberg (2008), Tetlock (2008), and Demers and Vega (2011) all examine the tone of firm-specific news items and find that the level of firm-specific news tone predicts future firm-specific earnings and returns. We add to this literature by showing that the cross industry news tone contains valuable information about firm fundamentals and future returns. Moreover, cross industry news provides more information inflow compared to firm news (news inflow of cross industry news is 7,436.69 per week while it is 2.42 for firm specific news). The overall abnormal return generated by cross industry news cannot be ignored and it serves a strong evidence that media news plays information intermediary that contributes to a better information environment.

Besides, CIS fully explore the profitability of cross industry return based portfolios, not vice versa. This finding serves strong evidence that media news contains information of lagged returns but also add on additional soft information to the market, consistent with the spirit of Tetlock (2008).

Third, we also contribute to the channels of information diffusion. Previous studies focus on the firm specific news traveling. For example, to use analyst coverage as proxy for information environment, Hong (2000) point out that bad news diffuses only gradually across the investing public. Similarly, Peress (2015) use news strike as an exogenous shock and demonstrate that the media influence the stock market by increasing the speed with which information diffuses across investors. Different from those literature, we look at cross industry news traveling and use news tone as direct measure of information. Based on our empirical design, we find direct evidence that analyst forecasts, institutional fund flows and media news are three channels help interpret cross industry information to the market.

The rest of the paper is organized as follows. Section 2 introduces the data and key variables used in this paper. Section 3 explains how to construct Cross Industry News Signal (CIS). Section 4 proposes CIS based trading strategy that can generate significant alpha and add investment value to existing risk factors. Section 5 shows the robustness of cross industry news information. Section 6 examines 3 channels of news traveling with Section 7 a brief conclusion.

2 Data and Key Variables

Our sample is collected from 5 major datasets. The news sample is from Thomson Reuters. Individual analysts' annual earnings forecasts and other related information are obtained from the I/B/E/S details file and institutional fund flow data is collected from EPRF databse. The data for firm financials and stock market variables are obtained from the

Compustat and CRSP databases.

We construct the news sample using all firm-specific news articles for all U.S. firms from Jan 1996 to Dec 2014. We require all news articles are novelty news which means it is the first time released or record by Thomson Reuters. We retrieve 11.63 million news stories from Reuters News Archive database. We then conduct textual analysis to read qualitative information of each news story according to the sentiment word list of Loughran (2011). Meanwhile, we use a variation of the approach in Hu (2004) to account for sentiment negation. If the word distance between a negation word (not, never, no, neither, nor, none, nt) and the sentiment word is no larger than 5, the positive or negative polarity of the word is changed to be the opposite of its original polarity. Following the literature (Tetlock (2008)), we measure news tone as the negative word ratio for each news story as:

$$\text{Tone} = \frac{\text{\# of negative word occurrences}}{\text{Total \# of Words in the news}}.$$

We then calculate firm specific news tone by averaging all news articles related to the firm following Huang et al., (2014) as:

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

where D stands for total number of firm specific news at time t.

Meanwhile, we find peer industry news for firm i as average news tone of peer firms within the same industry:

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

where k $\neq$ i and I stands for total number of firms of industry j excluding firm i. We then further define cross industry news of firm i as average of Industry News of those industries not covering firm i, namely:

$$\text{Cross Industry News}_{i,t} = \text{Industry News}_{j,t},$$

where $j \in \{1, 2..., N\}$ and N stands for total number of industry portfolios excluding the one firm i belongs. Indeed, to control for number of news effect (Fang and Peress (2009)), we also calculate the number of firm specific news, number of peer industry news and total number of cross industry news as additional controls of news effects.

On top of that, to examine whether cross industry news contains valuable information of firm performance, we follow Tetlock (2008) to use firm's standardized unexpected earnings (SUE) as proxy for firm fundamental. SUE is defined as Bernard (1989) and Thomas (1989):

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

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

where E_t is the firm's earnings in quarter t, and the trend and volatility of unexpected earnings (UE) are equal to the mean ($\overline{\text{UE}}$) and standard deviation (Std(UE)) of the firm's previous 20 quarters of unexpected earnings data, respectively. We also include control variables such as firm's size, B/M, turnover, three measures of recent stock returns and analyst dispersion. We define firm size (Log(Market Equity)) and B/M (Log(Book/Market

Equity)) at the end of the preceding calendar year, following Fama (1993). We compute turnover as the log of annual shares traded divided by shares outstanding (Log(Share Turnover)) at the end of the preceding calendar year. We also calculate analyst dispersion as the standard deviation of analysts' earnings forecasts within 3 to 30 days prior to the earnings announcement scaled by earnings volatility.

Besides, we follow Tetlock (2008) to calculate past returns based on a simple event study methodology. To align the estimation window, we choose the analysts' forecast announcement day or earnings annoucement day as the event day in accordance with dependent variable. Specifically, the benchmark return is calculated using the Fama-French three-factor model with an estimation window of [-252,-31] trading days before the event day. We also calculate the cumulative abnormal return on day -2 before the event day, denoted as $\text{CAR}_{t-2,t-2}$ and the cumulative abnormal return from the [-30,-3] trading day window before the event day, denoted as $\text{CAR}_{t-30,t-3}$. We further include the abnormal return from the estimation window, denoted as $\text{AR}_{t-251,t-31}$. In particular, $\text{AR}_{t-252,t-31}$ is related to the Jegadeesh and Titman (1993) return momentum effect, which is based on firms' relative returns over the previous calendar year excluding the most recent month. In addition, all the three past returns are presented in percentage.

In addtion, to follow Druz (2015), we add more firm characteristics and market conditions as control variables as following:

Market return is defined as the percent value-weighted market return for the period starting 5 days after an earnings announcement for the quarter t1 and ending 5 days prior to the earnings announcement for the quarter t. Momentum is defined as the firms buy-and-hold return over the prior 6 months. Illiquidity is defined as the absolute value of the stock return scaled by the product of volume and price. Leverage is defined as the long-term debt scaled by the sum of long-term debt and market capitalization. Institutional Ownership is defined as institutional share holdings scaled by shares outstanding. Monthly volatility is the monthly stock volatility computed from monthly return data over the previous 48 months and then classified into 10 quantiles. Besides, as standard control variables, we use the natural logarithm of market cap, as well as firm fixed effects, and year fixed effects.

Panel A of Table I provides summary statistics for earnings announcement related variables. For the average firm in our sample, there are on average 57,506.56 cross industry news, 1,237 peer industry news and 28.46 firm specific news 90 days before earnings announcement. The average industry news tone is 0.045 with individual industry news tone ranges from 0.038 to 0.052. The volatility of industry news is much smaller than firm specific news which suggests the little noise of industry portfolio. In the following analysis, to have fair comparisons among industry news tone effect, we normalize all news tone following Tetlock (2008).

< **Insert Table 1 here** >

Moreover, to support the argument that cross industry news deliver valuable information that results in cross industry reuturn predictability, we conduct Fama-Macbeth regression at weekly frequency. We use CIS (Cross Industry News Signal) as proxy for cross industry news (details have been explained in next section). Control variables include size, B/M, Leverage, Turnover, Return Volatility and # of firm specific news, industry news and total cross industry news. Panel B of Table I reports summary statistic for Fama-Macbeth regression at weekly frequency. It suggests that # of Cross Industry News is much larger than that of firm specific news - a potential large information is deliverred by cross industry news. In the following section, we try to explore this information by constructing CIS index for each individual firm.

3 Cross Industry News Signal (CIS)

To construct cross industry news information that may contribute to the future return, one needs a proxy of cross industry signal as the sorting variable. Unfortunately, due to complex industry interdependencies, different industries may have different effects on cross industry firms. In Table 2 of Rapach (2015), there are a sizable number of both positive and negative coefficient estimates, serving a direct evidence of complex economic linkages in the real world. In this paper, with an approach similar to that in Han *et al.* (2016), Rapach (2015) and Dashan Huang (2017), we extract information from multiple observables to obtain expected future stock returns, denoted by cross industry news signal (CIS). Our approach consists of three steps. In the first step, in each week t, we calculate news tone of each cross industry of firm j over the most recent week t-1 as Industry $\text{News}_{\bar{I_j},t-1}$. In sum, we use Fama-French 30 industries news tones as predictors to capture a firm's fundamental value. Indeed, for industry news tone that overlaps with industry classification of target firm, we set news tone of that industry as 0.

In the second step, in week t, for each individual firm j, we calculate out-of-sample forecast return based on cross industry news tones. To some extend, the general predictive regression model follows:

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

where $r_{i,t}$ is the week-t return on firm i in excess of the one-month Treasury bill return, N is the total number of industry portfolios and $\epsilon_{i,t+1}$ is a zero-mean disturbance term. To form CIS, we set Industry $\text{News}_{j,t}$ equal 0 when firm i belongs to industry j in above regression model. To include enough observations for model estimation, we require at least 260 weekly observations for each firm and set the initial estimation window as 208 weeks (4 years observation).

Moreover, to improve estimation and inference and avoid overfitting for the general predictive regression, we further employ the adaptive LASSO following Zou (2006) and Rapach (2015)). 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:

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

where Industry News$_{j,t}$ is the standardized news tone for industry j, $\hat{b}_i^* = (\hat{b}_{i,1}^*, ..., \hat{b}_{i,N}^*)'$ is the N-vector of adaptive LASSO estimates, λ_i is a nonnegative regularization parameters, and $\hat{w}_{i,j}$ is the weight corresponding to $|b_{i,j}|$ for j = 1, ..., N in the penalty term. Adaptive LASSO use L1-norm penalty to shrink the parameter estimates to prevent overfitting and hence, selecting most informative predictors. We then calculate cross industry news implied out-of-sample forecasted return for time t + 1 using adaptive lasso estimation results with information available at time t. This implied out-of-sample forecasted return is then defined as Cross Industry News signal:

$$\text{CIS}_{i,t} = \alpha_i + \sum_{j=1}^{N} E_t[b_{i,j,t+1}]\text{Industry News}_{\text{j,t}},$$

where $E_t[b_{i,j,t+1}]$ is the expected coefficient on the j^{th} industry and is defined as $\hat{b}_{i,j,t+1}$. CIS index is a real-time predictor of stock return and does not suffer from looking-forward biases.

Now, we are ready to construct CIS portfolio. In the following portfolio construction, we only sort stocks with negative CISs due to uninformativeness of positive CIS. Details have been discussed in the next section. At the end of each week, we sort all stocks into 10 value-weight portfolios by CISs, with the bottom quintile containing stocks with the lowest CIS and the top quintile containing stocks with the highest CIS. CIS portfolio is the zero-investment strategy that buys the top CIS portfolio and sells the bottom CIS quintile portfolio.

In the end of the section, we highlight the advantage of using CIS to rank stocks. According to Dashan Huang (2017), out-of-sample forecasted return based The signal combines multiple industry news tones to reflect the overall cross news information. Indeed, a collection of industry news have a strong and consistent impact on stock returns but it is weak with only one observable.

4 Empirical Results

4.1 Cross Industry News and Standard Unexpected Earnings

Our first set of analyses examines the link between cross industry news tones and actual earnings. We perform the following regression analysis:

$$\text{SUE}_{it} = \alpha_i + \sum_{j=1}^{N} \beta_j \text{Industry News}_{j,t-90,t-3} + \gamma^{`}\text{X} + \epsilon_{jt}, \text{where } I_k \neq I_j$$

where the dependent variable, SUE, measures each firm's standardized unexpected earnings following Bernard (1989). In this analysis, we start with US 1,295 firms for a period from 1996 to 2014, and we arrive at sample of 28,830 firm-quarter observations after losing observations in the process of merging with COMPUSTAT, CRSP, IBES, and the media data. Industry News$_{j,t-90,t-3}$ stands for the news information of cross industries within the period (t-90, t-3) relative to the earnings announcement day. Control variables include those firm specific news tones, peer industry news tones, # of firm specific news, # of peer industry news, # of cross industry news and those suggested by Tetlock (2008), including firms' lagged earnings (proxied by last quarter's SUE, lagSUE), Size, B/M, Turnover, three measures of recent stock returns ($\mathrm{AR}_{t-252,t-31}$, $\mathrm{CAR}_{t-30,t-3}$ and AR_{t-2}), analysts' earnings forecast revisions (Forecast Revision), analysts' forecast dispersion (Analyst Dispersion). Besides, we further control other variables documented in related literatures (Jegadeesh (2004) and Druz (2015), among others), including dummy variable of news coverage ($I_{newscoverage}$), Consensus Forecast, Management Forecast, Earnings Surprise, Return Volatility, Market Return, Institutional Ownership, Leverage, Momentum, Illiquidity and Overconfidence. Based on this setting, if cross industry news tone show significant prediction power on actual earnings, we may expect cross industry news also captures valuable information of firm fundamentals to investors.

< **Insert Table 2 here** >

Table 2 presents the panel regression results, with standard errors clustered by firms. In Panel A, we only include one cross industry news tone in each regression model. The first 3 columns show estimation coefficients, T-value and adjusted R^2 for univariate test. The middle 3 columns show corresponding results that follow Tetlock (2008) setting. In the last 3 columns, we add on all other control variables to serve a stronger tests on cross industry news information. Indeed, results are quite consistent across different settings. Most cross industry news tone negatively predicts firm's earnings with significance at 1% level. Only news tone of Coal industry positively predicts SUE. This is consistent with the argument that Coal industry serves as most important supply chain so a negative shock of Coal industry reduces the cost of supply side and it positively affect the earnings of industries located in later stages of production processes. While according to Hong and Valkanov (2007), a better way to test the cross industry effect is to control all cross industry variables into regression. The benefit of doing this is that since industry returns are contemporaneously correlated, by including all variables, we do not worry about issues related to omitted variables. In Hong and Valkanov (2007), the cost of doing this is that the standard errors on estimates will be larger due to a limited number of observations. In our case, we study at firm level, which helps to handle this issue. In Panel B, we reports results by including all cross industry news into one regression. The results changes a lot due to interactions of cross industry news information. Indeed, many news tone of industries become insignificant or change their prediction signs. Those industries remain strong prediction power on other firms' earnings include Food, Beer, Smoke, Books,Hlth, ElcEq, Autos, Mines, Paper and Trans. Different from lagged return predition documented in Rapach (2015), media news suggests those demand chain play an important role in affecting next period earnings of other industries. While core industries, like Coal and Oil does not show significant power under the new setting. We explain it that might because we use most recent 20 years observations due to limitation of media news data and it is possible that in the recent years, the market is mainly driven by demand side instead of supply side. On top of that, loading on those industry news tones exhibit substantially positive and negative predictions on SUE, suggesting a complex industry interdependencies that have bullish implications for some industries and bearish implications for others. Under adaptive LASSO, ElcEq, Autos, Coal and Util are the most informative predictors for other industry earnings. Consistent with Rapach (2015), a negative news tone of Coal industry serves as good news to other industries due to a decreasing cost from upper stream industry. Meanwhile, ElcEq, Auto and Util are 3 new technique related industries to reflect a more complicated industry linkage, such as technique spill-over effect.

4.2 Cross Industry News and Stock Returns

Having established that cross industry news can predict firms' fundamentals, we now examine whether they provide novel information not already represented in stock market prices. As we have observed in previous section, firms in different industries may react differently to other industry news. In this case, we calculate CIS for each individual form and take CIS as proxy of cross industry news related information shock. Detail construction of CIS is shown in section 3.

4.2.1 CIS Value under Time Series Setting

After that, we first examine the market's apparently sluggish reaction to CIS shocks in 62 weeks before and after the week that shows extreme CIS. Figure I graphs a firm's abnormal returns around the extreme CIS event. We use stock return minus equal weight market return to estimate abnormal returns. We label all stocks with out-of-sample forecasted return in the top (bottom) decile as high (low) CIS group. Figure I shows market seems asymmetrically react positive and negative CIS shocks. Investors tend to overreact to positive CIS shocks on the event day and then abnormal return lowers to 0 or even negative. This may suggests that positive CIS is not informative after event and investors tend to overreact to positive shocks on the event day. While it is not the case for negative CIS. We find abnormal return after extreme negative CIS signal continues a significant negative pattern. Indeed, it has not been corrected by the market for quite a long period. The persistence of prediction serves as an evidence of slow information diffusion across industries, consistent with Hong and Valkanov (2007).

< Insert Figure 1 here >

Panel B of Figure 1 shows results based on out-of-sample forecast return using lagged cross industry return. The pattern is quite similar to CIS, which may suggest lagged return and CIS share a similar information set. Indeed, positive signal disappears after the even while negative shock continues. Consistent with property of firm specific news traveling Hong (2000), we document negative news travels slowly across industries.

On top of that, Figure 1 shows pre-event abnormal returns that is significantly different from 0. This pre-event premium is not completely due to informed trading. Different from firm specific news, cross industry news comes in for each trading trading. Prior to the extreme event (bottom decile of CIS), CIS shows negative signals to the market. Figure 2 plots the CIS signals around the same event day as Figure 1 shows. The signal seems quite efficient with narrow bands of 95% interval. It also suggests that CIS has some continuation with a significant negative signal even around 40 weeks ago. This is good for trading as it reduces the turnover a lot compared to firm specific news.

< Insert Figure 2 here >

4.2.2 CIS Value under Cross Sectional Setting

Above analysis shows valuable cross industry news at time series aspect. After that, we conduct Fama-MacBeth regression to test cross sectionally premium of CIS. The advantage of Fama-MacBeth is that one can control for other firm characteristics, which may contain information in the variables of interest. Accordingly, we choose Lagged Return, Size, B/M, Leverage, Turnover, Return Volatility, Firm News, Industry News, # of Firm News, # of Industry news and # of Cross Industry News as control variables. Indeed, since positive signal is not informative to market, we only consider stocks that shows negative CIS.

Regression results have been shown in Table 3. We have 3 different sample periods and 3 groups of control variables. The first 3 columns use the whole sample period, namely 2000-2014 (we take 1996 to 1999 as initial estimation window), the middle 3 columns present results of 2000-2007 and the last 3 columns show results of 2008-2014. Overall, the results are quite consistent across different settings. CIS shows strong cross sectional premium of stock returns. For the overall sample period, 1% increase in CIS, stock return tend to increase by 2.25% given other situations fixed. Overall, the evidence suggests that cross industry news is not well priced by the current market.

< **Insert Table 3 here** >

The lingering difference of cross sectional premium of CIS suggests that a simple trading strategy could earn positive risk-adjusted profits. In this section, we explore this possibility, focusing on the apparent underreaction to negative CIS. Specifically, at the close of each trading week, we form two equal-weighted portfolios based on the content of each firm's CIS during the prior trading week. We define the lowest decile of CIS as short leg while the highest decile of CIS as long let. Indeed, we only use non-positive CIS to sort stocks due to uninformativeness of positive CIS. We then hold both the long and short portfolios for 1 full trading week and rebalance at the end of the next trading week. Ignoring trading costs, the cumulative raw returns of this long-short strategy would be 9.48% per year. Figure 3 shows cumulative return of CIS and equal weight stock returns of whole samples. It seems that our CIS strategy performs extremely well during the recession periods while equal weight return drops a lot. This suggests a valuable information of cross industry news.

< **Insert Figure 3 here** >

Table 4 shows the risk-adjusted weekly returns from this weekly news-based trading strategy for three different time periods (2000 to 2014, 2000 to 2008, and 2009 to 2014). We use the Fama (1993) and Carhart (1997) models to adjust the trading strategy returns for the returns of contemporaneous market, size, book-to-market, and momentum factors.

The first column of each sample period reports the results with the market risk benchmark, the middle column reports the results of Fama-French benchmark, whereas the last column uses the Carhart benchmark. We compute all coefficient standard errors using the White (1980) heteroskedasticity-consistent covariance matrix. Consistent with Table 3, Table 4 shows that the weekly CIS-based trading strategy would earn reasonable good risk-adjusted returns in a frictionless world with no trading costs or price impact. Specifically, the average excess return (Fama-French alpha) from CIS-based trading would be 21 bps per week from 2000 to 2008 and 22 bps per week from 2009 to 2014. The increased benefit may suggest a closer connections among different industries in the recent years. Overall, using any return benchmark, the alpha from the trading strategy is highly significant in all three sample periods.

< **Insert Table 4 here** >

For the 15 years between 2000 and 2014, Figure 4 depicts the distribution of the average daily abnormal returns for the CIS-based trading strategy. In the median year, the strategy's abnormal return is 4.4 bps per day. In 13 out of 15 years, the CIS-based strategy earns positive abnormal returns. Thus, we can reject the null hypothesis that yearly CIS-based strategy returns follow the binomial distribution with an equal likelihood of positive and negative returns (p-value ¡ 0.0005). There is only 1 year out of 15 in which the strategy lost more than 5 bps per day. By contrast, in 6 out of 15 years, the strategy gained more than 5 bps per day. This analysis suggests that the CIS-based trading strategy is not susceptible to catastrophic risks that second moments of returns may fail to capture.

< **Insert Figure 4 here** >

Further more, we estimate the impact of reasonable transaction costs on the trading strategy's profitability. To judge the sensitivity of profits to trading costs, we recalculate the trading strategy returns under the assumption that a trader must incur a round-trip transaction cost of between zero and 10 bps. Table 5 displays the abnormal and raw annualized CIS-based strategy returns under these cost assumptions. We also show firm specific news based portfolio performance as a benchmark. From the evidence in Table 5, we see that the simple firm specific news-based trading strategy is no longer profitable after accounting for reasonable levels of transaction costs, for example, 10 bps. While CIS-based portfolio survives even under 10 bps transaction cost. This suggests that turnover of cross-industry-news based portfolio is much smaller than that of firm specific news.

< Insert Table 5 here >

Above analysis suggests the cross industry news contains valuable information to predict future stock returns and the benefit is reasonable good. In this section, we further study how cross industry news decays as time goes by. We recalculate the portfolio returns under the assumption of different prediction horizons, namely 1 to 10 weeks after CIS signal. To be more specific, we rebalance the portfolio at the end of current week using CIS 10 weeks ago. Table 6 reports results under different forecast horizons. Consistent with our expectation, cross industry news has a long and persistent effect on future stock returns than firm specific news. The risk adjusted return of CIS remains 10.85% annulized return for 10-week ahead signal while firm specific news lost its significance after 4 weeks and its raw return drops more than 50% in the second week. This further confirms the argument that news travels slowly across industries.

< Insert Table 6 here >

5 Robustness Check

In this section, we further consider some alternative explanations that could drive our results and show that cross industry news remains power after controlling for various market effects and is robust to alternative research designs.

5.1 Impact of Overlapped Information

In this section, we consider overlapped information between cross industry news and alternative information source, including firm specific news, peer industry news and lagged returns of 30 industries. The predictability of cross industry news could be driven by the overlapped information with those related variables, hence, the cross industry return predictability of media news is not surprising. To investigate this question, we build on Table 4 by adding 3 additional factors.

< Insert Table 7 here >

Table 7 reports results of alternative information adjusted alpha of CIS strategy. Similar to Table 4, we divide sample periods into 3 periods, namely 2000 - 2014, 2000 - 2008 and 2009 - 2014. For the first column of each sample period, we add on the portfolio return based on peer industry news. For the second column, we add on the portfolio return based on firm specific news. In the 3rd column, we add on portfolio return based on the out-of-sample forecasted return using lagged returns, including firm lagged return, industry lagged return and cross industry lagged return. Indeed, we find Cross Return portfolio shows very strong explanation power on the CIS portfolio. but alpha remains positive with significance level at 1%, suggesting that cross industry news contains additional soft information that is not priced by the market, consistent with Tetlock (2008). In addition, we also run Cross Return portfolio on CIS and find the alpha of Cross Return is fully explained by the CIS strategy. This serves strong evidence that CIS delivers the information of lagged industry return and also new information to the market. Overall, Table 7 suggests that CIS survives the tests of alternative information sources. It not only explains the cross industry return predictability but also contribute new information to the market.

5.2 Impact of Investor Sentiment

News tone could be mixed by soft information and journalists' sentiment. Although we have provide evidence that cross industry news predicts future earnings, it is possible that it reflects overall market sentiment that contributes to the return predictability. In this section, we use Baker (2006) sentiment and Huang (2014) PLS sentiment index to stand for the aggregate investor sentiment in the stock market and take market wide news tone to stand for the market wide news "sentiment". We then define a period as a high sentiment period if the sentiment index is above the median of the whole sample period and a low sentiment period otherwise. We then evaluate the profitability of CIS portfolio performance over high and low sentiment periods, respectively.

Table 8 shows that CIS strategy is associated with investor sentiment, and the profits is stronger in high sentiment periods than that in low sentiment periods. During high sentiment periods, the most optimistic views tend to overly optimistic and stocks are more likely overpriced. In contrast, during low sentiment periods, the most optimistic views tend to be closer to those of rational investors and stocks are more likely to be correctly priced. As a result, mispricing is more likely during high sentiment periods, consistent with the spirit of Stambaugh (2012). Different from investor sentiment, we find the profits concentrate in the low market wide news tone periods, which is consistent with previous literature that negative news is more informative than the positive news.

Overall, this section shows that CIS is only partially explained by short-sale impediments in the stock market, due to potential institutional constraints, arbitrage risk, behavioral biases of traders, and trading costs.

< Insert Table 8 here >

5.3 Impact of Macro Environment

Cross industry news could also be proxy for macro news which might be related to macro environment. We then test CIS sensitivity to the market wide uncertainty index, including VIX, EPU (Baker (2015)) and market wide news dispersion. We define market wide news dispersion following Dzieliski (2015). Similar to analysis of sentiment effect, we define a period as a high uncertainty period if the uncertainty index is above the median of the whole sample period and a low uncertainty period otherwise. We then evaluate the profitability of CIS portfolio performance over high and low uncertainty periods, respectively.

Table 9 suggests that CIS concentrates in the high uncertainty periods. After controlling all variables, alpha in high VIX periods outperforms the low VIX periods by 5% annualized return. Indeed, CIS also generates positive alpha during the low VIX periods. Again, VIX only explain part of the CIS performance. In terms of EPU, CIS seems insensitive to economic policy uncertainty with little difference of alpha between high and low EPU periods. When it comes to news dispersion, we find the result is still robust. After controlling for firm news and peer industry news, alpha becomes much smaller for the low news dispersion periods. This is consistent with our expectation. For cross industry news informative to the market, it should deliver something new. If disagreement is low among firms, cross industry news overlaps a lot with peer industry news and firm specific news, hence swallowing the cross industry news information. Indeed, CIS concentrates in the high news dispersion periods serves an additional support to the cross industry information diffusion story.

< Insert Table 9 here >

5.4 Impact of Information Environment

Given CIS contains soft information to investors, it should be more valuable to those firms with poor information environment. According to Fama (2015), smaller firms tend to have higher mispricing. This fact raises the question of whether cross industry news concentrates heavily in small firms. This argument also applies to high volatility stocks, low analyst coverage stocks and illiquidity stocks. As a result, we double sort CIS with other information environment proxy, including size, volatility, illiquidity, analyst coverage and analyst dispersion. We then plot portfolio cumulative return in Figure 5.

< Insert Figure 5 here >

Consistent with our expectation, CIS strategy performs better within groups under poor information environment, suggesting that the cross industry news information is more valuable to those stocks. For example, CIS reveals valuable information to both liquid and illiquid stocks while it generates much higher cumulative return using illiquid stocks. In this case, illiquid stocks incorporate cross industry information with a delayed timer comparing to the liquid stocks hence resulting in lead-lag among these two groups of stocks.

6 Channel of Cross Industry News Travelling

Different from firm specific news, cross industry news is not easy to understand. In section III, we have observed that cross industry news show different predictions on firm fundamentals. For firms in the same industry, they may react to the same cross industry substantially different. In this case, without information intermediary that help interpret cross industry news to the market, it is difficult for news to travel across industries. In this section, We try to understand which type of market participants help interpret cross industry news to the market. We think of three main channels, namely analyst earnings forecast, institutional fund flows and media news.

6.1 Cross Industry News and Analysts Forecast Behavior

Analysts are usually regarded as sophisticated investors and also serves as information intermediary of financial market. If cross industry news contains valuable information about firm fundamentals, analysts should incorporate this information into their earnings forecasts. In this way, they interpret other industry information to the market and improve a firm's information environment. In this section, we examine whether cross industry news affects analysts' forecast behavior, such as forecast revisions and forecast improvement. To answer this question, we perform the following regression analysis:

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

where Y_{ijt} stands for Annual Forecast Revision$_{ijt}$ and ForecastImprovement$_{ijt}$. We define forecast revision as the absolute change of analyst forecasts scaled by stock price in the end of last year and define forecast improvement as the current forecast accuracy minus previous forecast accuracy for the same earnings forecast period. We define forecast accuracy as the minus absolute value of difference between actual earnings and analyst forecast. X denotes other explanatory variables which are defined in Appendix. Table 10 presents the panel regression results. The first 3 columns report results of forecast revision and last 3 columns show results of forecast improvement. Consistent with our expectation, analysts tend to adjust their forecast revisions when average cross news tone is high and the revised forecast tend to have a higher forecast accuracy than the previous forecast. In other words, Table 10 provide direct evidence that anlaysts incorporate cross industry news into their earnings forecast and improves their forecast accuracy. In terms of economic significance, the result in last column suggests that a one standard deviation of AverageCrossNewsTone$_{t-90,t-3}$ is associated with a 0.019% improvement in forecast accuracy. Overall, we provide evidence that analysts serve as information intermediary that interpret cross industry news to the market.

< Insert Table 10 here >

6.2 Cross Industry News and Institutional Investors

Under the sophisticated institutions hypothesis, institutions should explore valuable cross industry news to gain abnormal returns. In this case, we expect their fund flow should reflect cross industry news in a way they interpret the news. To avoid noise measure of institutional investors' behavior, we only use active institutional fund flow at industry level. We then study how cross industry news affect active institutional investors' fund flow. Fund flow data is collected from EPRF with industry labeled as GIC code. While in our study, to align with the main dataset from WRDS, we use SIC code to classify industries following Fama-French. In the first step, we map GIC classifications into Fama French 30 industries and then use cross industry news tone to predict subsequent industry fund flows. The regression model follows:

$$\text{Industry Fund Flow}_{kt} = \alpha + \sum_{j=1}^{N} \beta_j \text{Industry News Tone}_{\mathrm{j,t-1}} + \sum_{m=1}^{N} \beta_m \text{Industry Fund Flow}_{\mathrm{j,t-1}} + \epsilon_{kt},$$

where Industry Fund Flow stands for weekly active institutional fund flow for different industries according to GIC. We map this industry classification to Fama French 30 industries and take cross industry news tones as variables of interests.

Table 11 reports summary results of cross industry news effects on institutional investors' fund flow. The first 3 columns in Table 12 count the number of significant industry vraiables that is not overlapped with industry classification of Fund Flow. Column 4 to 7 reports adjusted R^2 that includes lagged industry fund flow alone, lagged industry news tone alone, industry fund flow and industry news tone, all cross industry news tones respectively. Last column reports F statistics using all cross industry news tones. Overall, we find cross industry news significantly affect active institutional fund flows. On average, there are 2 cross industry news tones show significance at 1% level that predict next period fund flows and there are additional 4 industries show significance at 5% level. Importantly, the adjusted R^2 by using only cross industry news tones is on average 4.82 times of using only industry fund flow and industry news tone (excluding the cases of other industry and Mixed industry, the average adjusted R^2 is 1.45% and 6.98% respectively). This remarkable result suggests institutions tend to explore cross industry news for their asset allocations, hence contributing to the news traveling across different industries.

< **Insert Table 11 here** >

6.3 Cross Industry News and Firm Specific News

Recently, Ying Wang and Zhu (2016) provide empirical evidence of news momentum, namely, the arrival of news can be predicted by historical news. This argument extends news traveling by studying whether cross industry news contributes to the firm specific news hence driving the return predictability of cross industry news. To test this hypothesis, we employ Fama-MacBeth regression running firms specific news tone on lagged CIS by controlling other firm characteristics following Table 3.

Table 13 reports empirical results. The first 3 columns use the whole sample period, namely 2000-2014, the middle 3 columns present results of 2000-2007 and the last 3 columns show results of 2008-2014. Overall, CIS consistently predicts next weeks' firm specific news tone even controlling for lagged firm return. In terms of economic significance, 1% increase in CIS, firm specific news tone tend to decrease by 7.54% given other situations fixed. This evidence further suggests media news serves an alternative channel that helps cross industry information diffusion.

< **Insert Table 13 here** >

7 Conclusion

In this paper, we provide direct evidence that cross industry news contains valuable information about firm fundamentals and this information has a longer forecast horizon on stock returns compared to firm specific news. To some extend, a long-short portfolio based on CIS generates an annulized risk adjusted return 10.85% after 10 weeks of the signal. Cross industry news also contributes to a better information environment and is more valuable to small stocks, illiquid stocks, and stocks with high return volatility. Analyst forecasts, institutional fund flows and media news might be the channels that interpret cross industry news to the market.

References

Ahern, K. R. and Sosyura, D. (2014). Journal of finance. *Quantitative Finance*, **69**.

Badrinath, J. R. K., S.G. and Noe, T. H. (1995). Of shepherds, sheep and the cross-autocorrelations in equity returns. *Review of Financial Studies*, (8), 401–430.

Baker, . W. J., M. (2006). Investor sentiment and the cross section of stock returns. *Journal of Finance*, **61** (4), 1645–1680.

Baker, B. N. . D. S. J., S. R. (2015). Measuring economic policy uncertainty. *NBER Working Paper*.

Barber, B. M. and Odeani, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, **81**, 785 – 818.

Bernard, . T. J. K., V. L. (1989). Media coverage and the crosssection of stock returns. *Journal of Accounting research*, pp. 1–36.

Boudoukh, M. P. R., Jacob and Whitelaw, R. F. (). *A tale of three schools: Insights on autocorrelations of short-horizon stock returns.*

Brennan, N. J., Michael J. and Swaminathan, B. (1993). Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies*, (6), 779–824.

Bushee, C. J. E. G. W. . H. S. J., B. 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.

Dashan Huang, G. Z., Huacheng Zhang (2017). Twin momentum. *Working Paper*.

Davis, P. J. M. . S. L. M., A. K. (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?s. *Working Paper*.

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

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

Engelberg, J. (2008). Costly information processing: Evidence from earnings announcements. *Working Paper*.

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

Fama, F. K., E.F. (2015). A five-factor pricing model. *Journal of Financial Economics*, **116**, 1–22.

Fang, L. and Peress, J. (2009). Credit expansion and neglected crash risk. *Working Paper*.

Gervais, R. K., S. and Mingelgrin, D. H. (2001). The high-volume return premium. *Journal of Finance*, **56**, 877–919.

Han, Y., Zhou, G. and Zhu, Y. (2016). A trend factor: Any economic gains from using information over investment horizons? *Journal of Financial Economics*, **122** (2), 352–375.

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

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

Hong, W. T., H. and Valkanov, R. (2007). Do industries lead stock markets? *Journal of Financial Economics*, **83**, 367396.

Hou, L. P., K. and Xiong, W. (2009). A tale of two anomalies: The implications of investor attention for price and earnings momentum. *Working Paper*.

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

Huang, J. F. T. J. . Z. G., D. (2014). Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies*.

Jegadeesh, K. J. K. S. D. . L. C., N. (2004). Analyzing the analysts: When do recommendations add value? *Journal of Finance*, **59** (3), 1083–1124.

Jegadeesh, N. and Titman, S. (1995). Overreaction, delayed reaction and contrarian profits. *Review of Financial Studies*, (8), 973–993.

Lo, A. and MacKinlay, C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, (3), 175–206.

Loh, R. (2010). *Investor inattention and the underreaction to stock recommendation*, vol. 39.

Loughran, . M. B., T. (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.

Peress, J. (2015). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *The Journal of Finance*, **69** (5), 2007–2043.

Rapach, S. J. T. J. . Z. G., D. (2015). Industry interdependencies and cross-industry return predictability. *Working Paper*.

Stambaugh, Y. J. Y. Y., R.F. (2012). The short of it: investor sentiment and anomalies. *Journal of Financial Economics*, **104**, 288302.

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.

Tetlock, S.-T. M., P. C. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, **63** (3), 1437–1467.

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

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

Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of Financial Economics*, **116**.

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

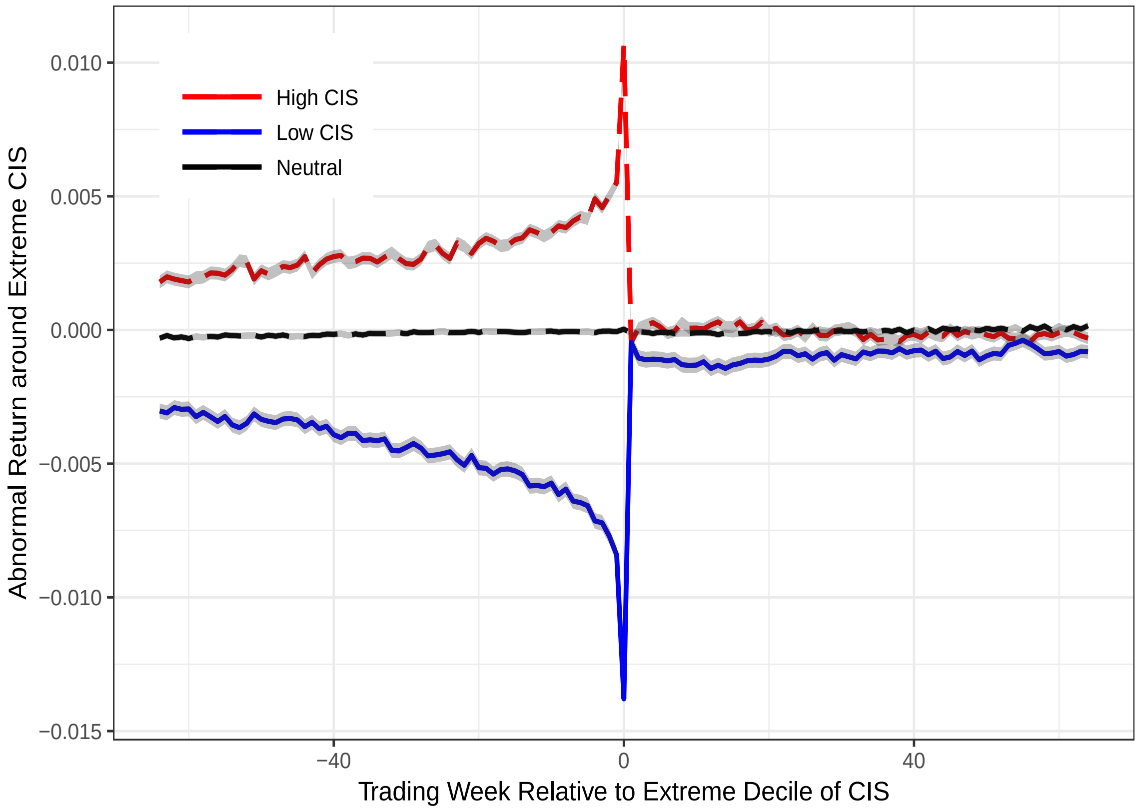
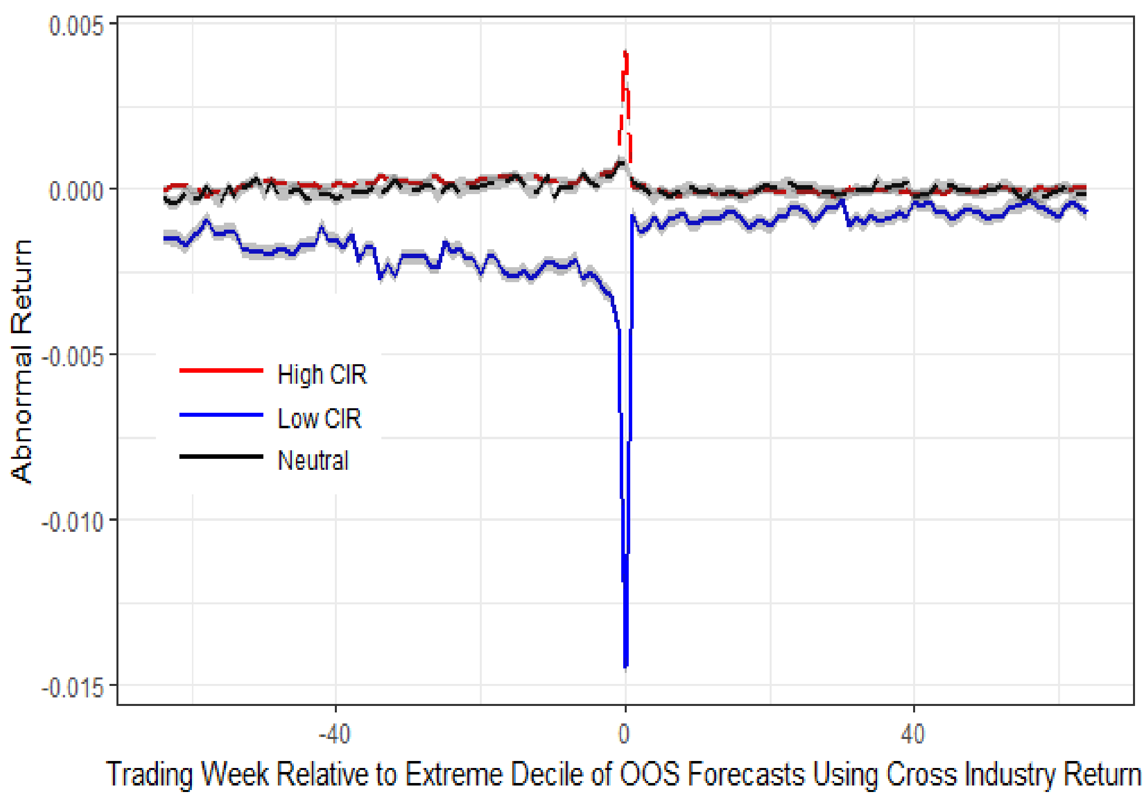


Figure 1: **Abnormal Return around Extreme Decile of CIS** In this figure, we graph a firm's abnormal returns from 64 trading weeks preceding a top (bottom) decile of out-of-sample forecasted return based on CIS (cross industry news signal) and CIR (cross industry return singal) respectively. All news stories come from Thomson Reuters between 1996 and 2014 inclusive. The out-of-sample period is 2000-2014. We calculate weekly abnormal return as stock return minus equal weight market returns. We separately examine the market's response to positive and negative signals. We also compute 95 confidence interval for both positive and negative shocks represented by the gray area.

Figure 1 (continued)



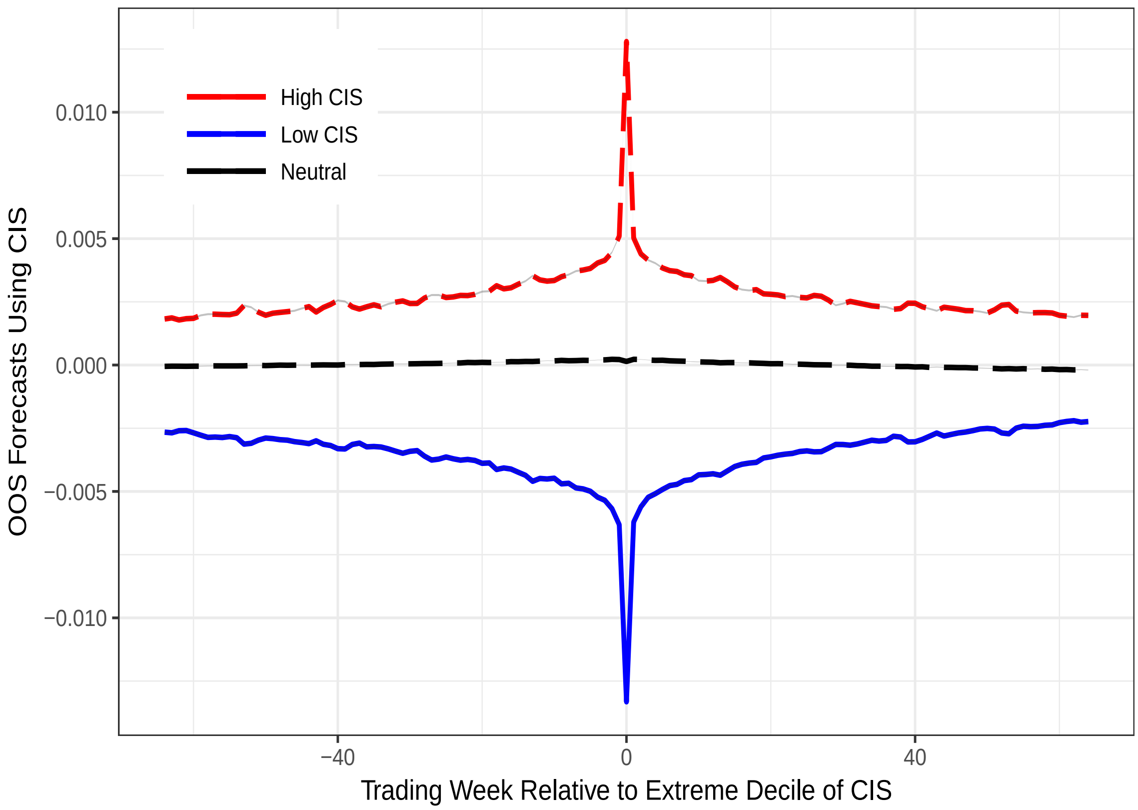
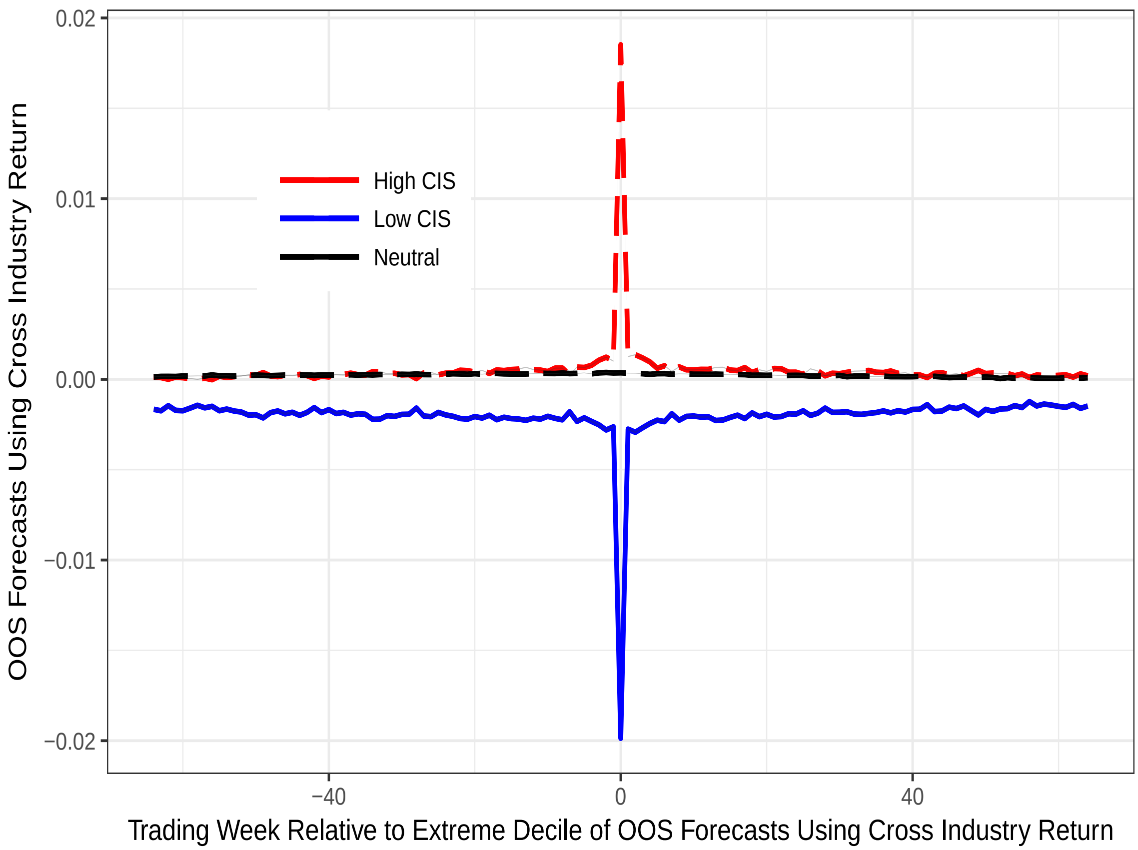


Figure 2: **Cross Industry News Signal (CIS) around Extreme Event** In this figure, we plot a firms' CIS (CIR) from 64 trading weeks preceding a top (bottom) decile of out-of-sample forecasted return based on CIS (CIR). All news stories come from Thomson Reuters between 1996 and 2014 inclusive. The out-of-sample period is 2000-2014. We also compute 95 confidence interval for both positive and negative signals represented by the gray area.

Figure 2 (continued)



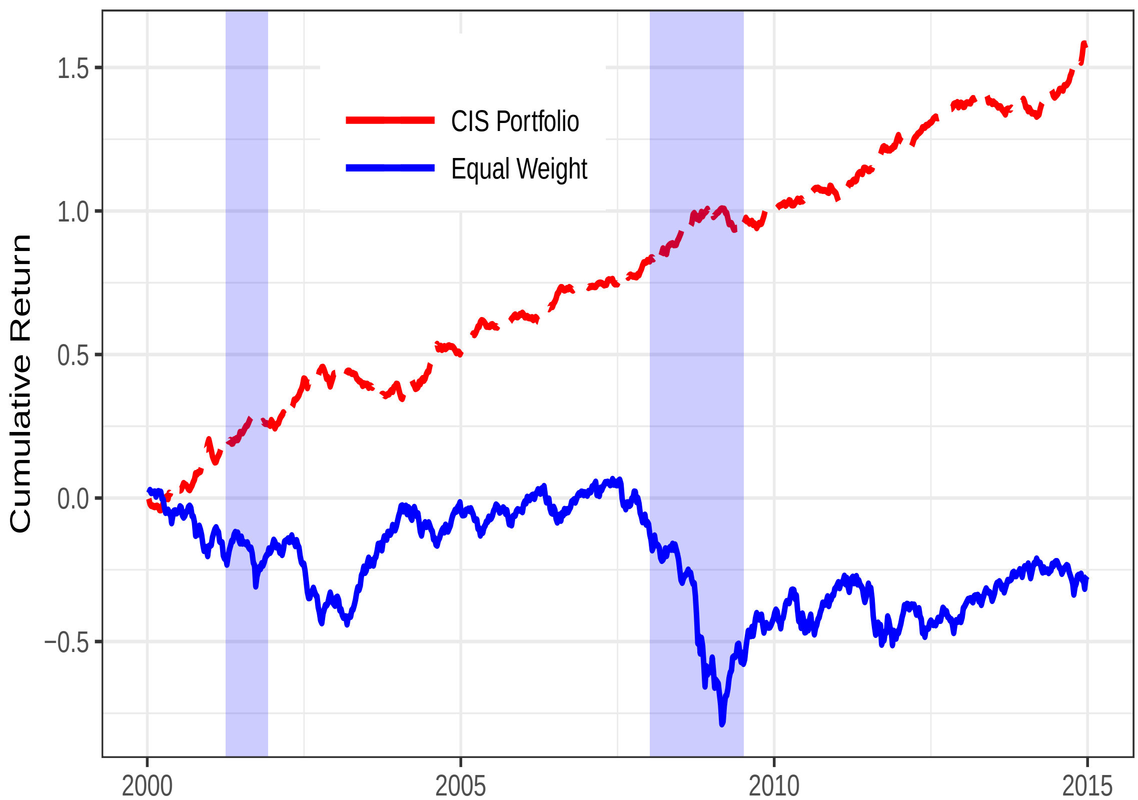


Figure 3: **Cumulative Returns of Cross-Industry-News (CIS) portfolio.** This figure shows the cumulative returns of investment strategies based on nonpositive CIS in prior trading week. The sample period spans from 1996-01-03 to 2014-12-31. We form two equal-weighted portfolios based on the signal of each firm's cross industry news from Thomson Reuters during the prior trading week. We label all stocks with CIS in the top (bottom) decile as long (short) leg. We hold both the long and short portfolios for 1 week and rebalance at the close price of next week.

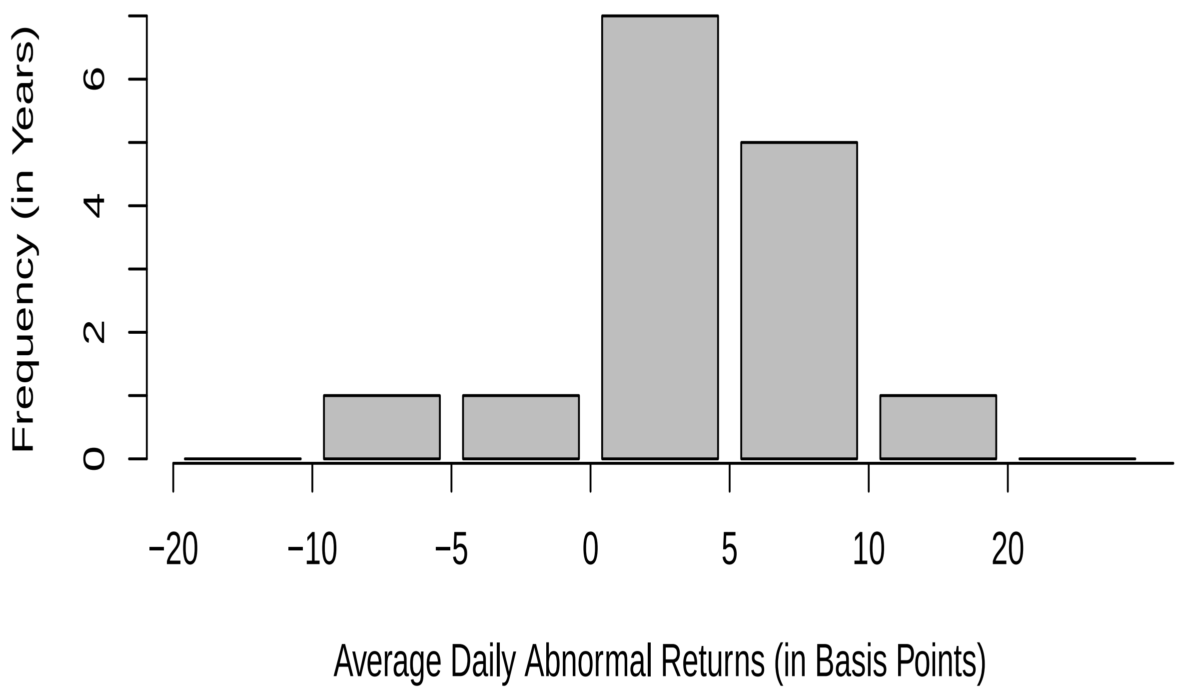


Figure 4: **Distribution of daily abnormal returns for the cross-industry-news (CIS) portfolio.** For the 15 years from 2000 to 2014, the figure depicts the distribution of the average daily abnormal returns of the cross-industry-news-based trading strategy described below. Each frequency bin encompasses a certain bps range of abnormal returns described by the two numbers adjacent to the bin – for example, the frequency of the leftmost return bin is the number of years in which the trading strategy's average daily abnormal return is between -20bps and -10bps. We assemble the portfolio for the nonpositive CIS based trading strategy at the close of each week. We form two equal-weighted portfolios based on our cross-industry-news signal (CIS) during the prior trading week. We label all stocks with out-of-sample forecasted return in the top (bottom) decile as long (short) leg. We hold both the long and short portfolios for 1 week and rebalance it at the close price of next week. To adjust the returns for risk, we calculate daily abnormal return as stock returns minus S&P500 index returns.

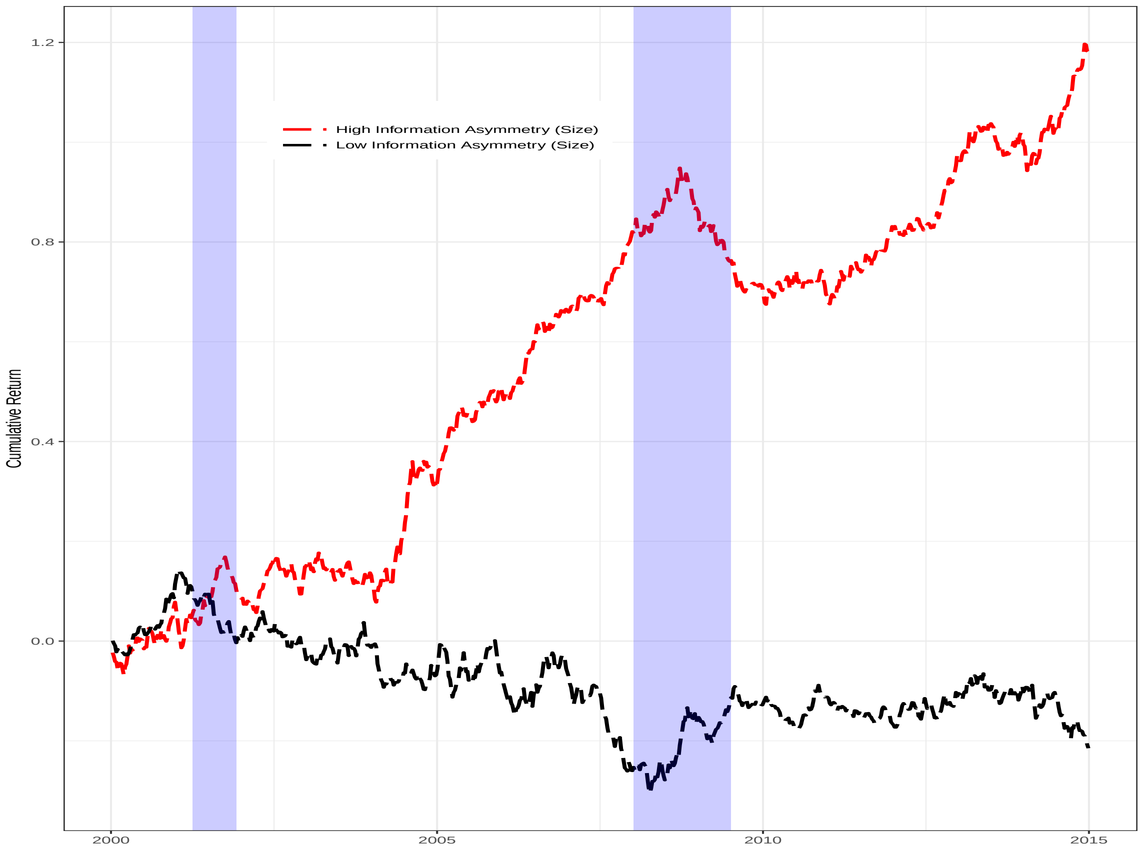
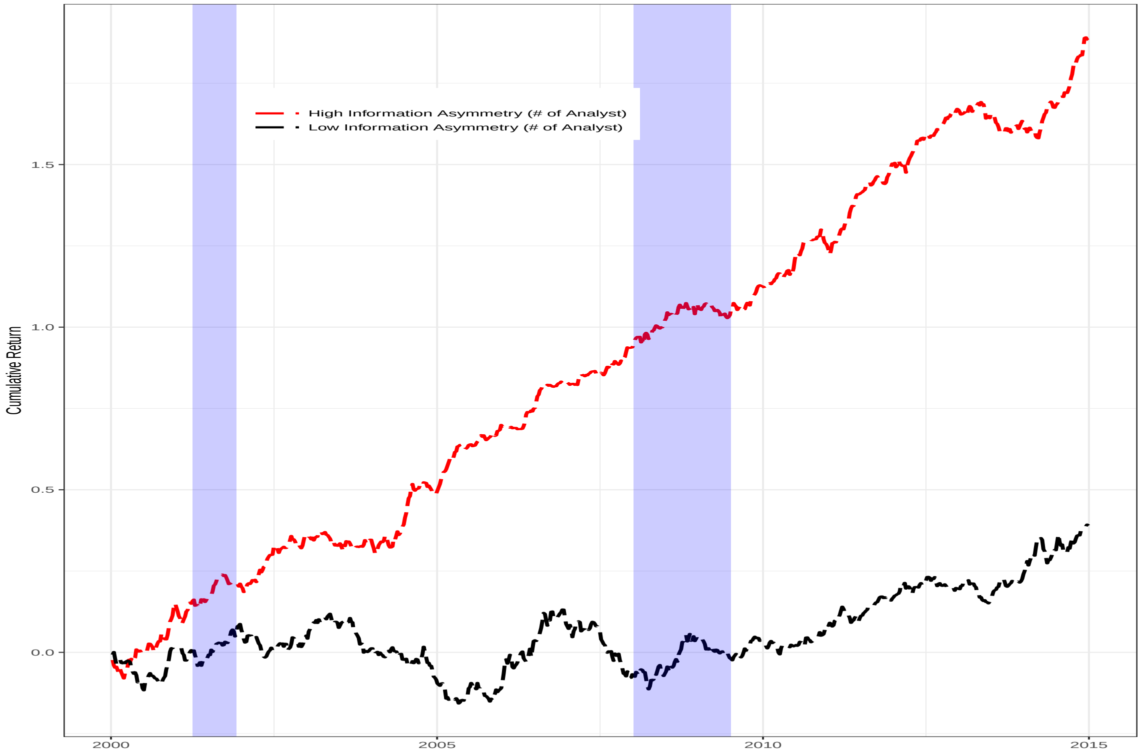


Figure 5: **CIS Based Portfolio Performance with High and Low Information Asymmetry Stocks** This figure shows the cumulative returns of investment strategies based on CIS in prior trading week within those high and low information asymmetry stocks. The information asymmetry proxy includes size, analyst coverage, analyst dispersion, liquidity and return volatility. The out-of-sample period spans from 2000-01-07 to 2014-12-26. We label all stocks with CIS in the top (bottom) decile as long (short) leg. We hold both the long and short portfolios for 1 week and rebalance it at the close price of next week.

Figure 5 (continued)



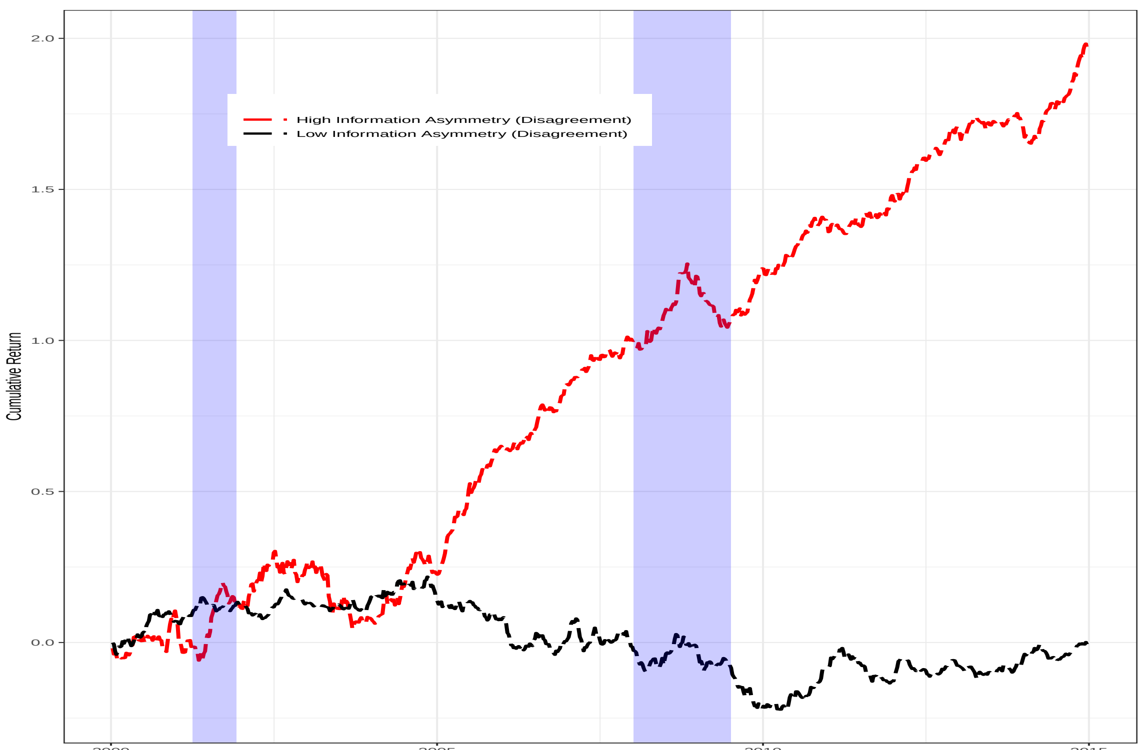
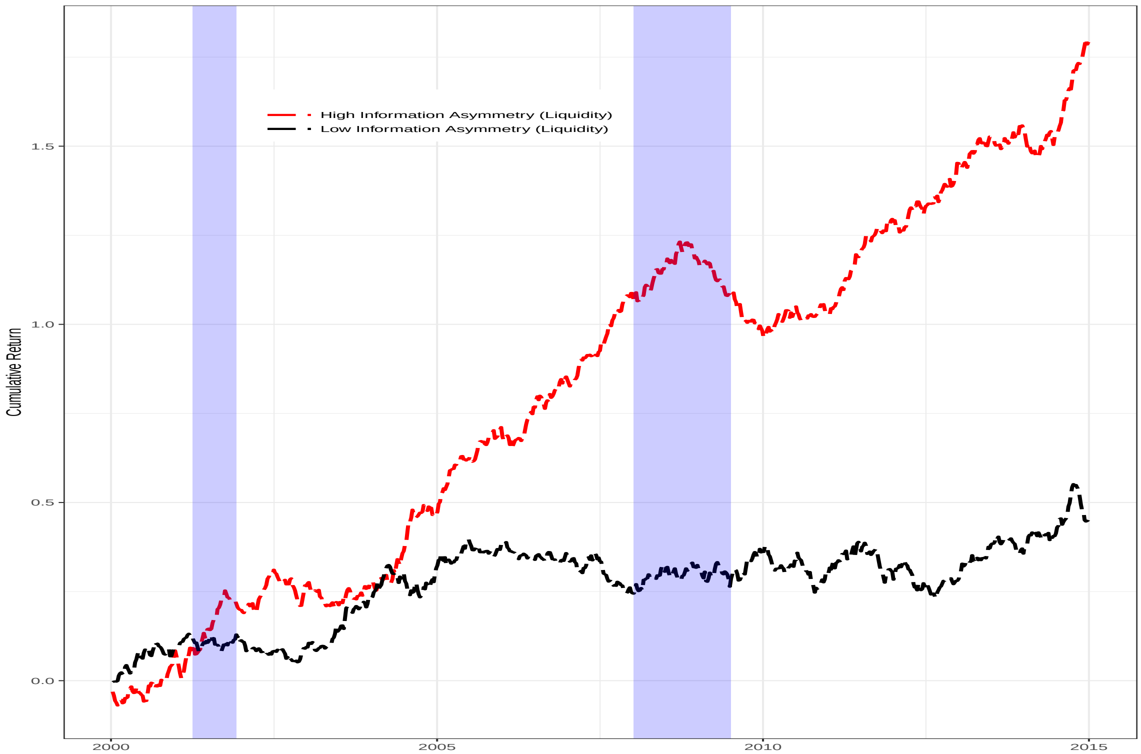


Figure 5 (continued)



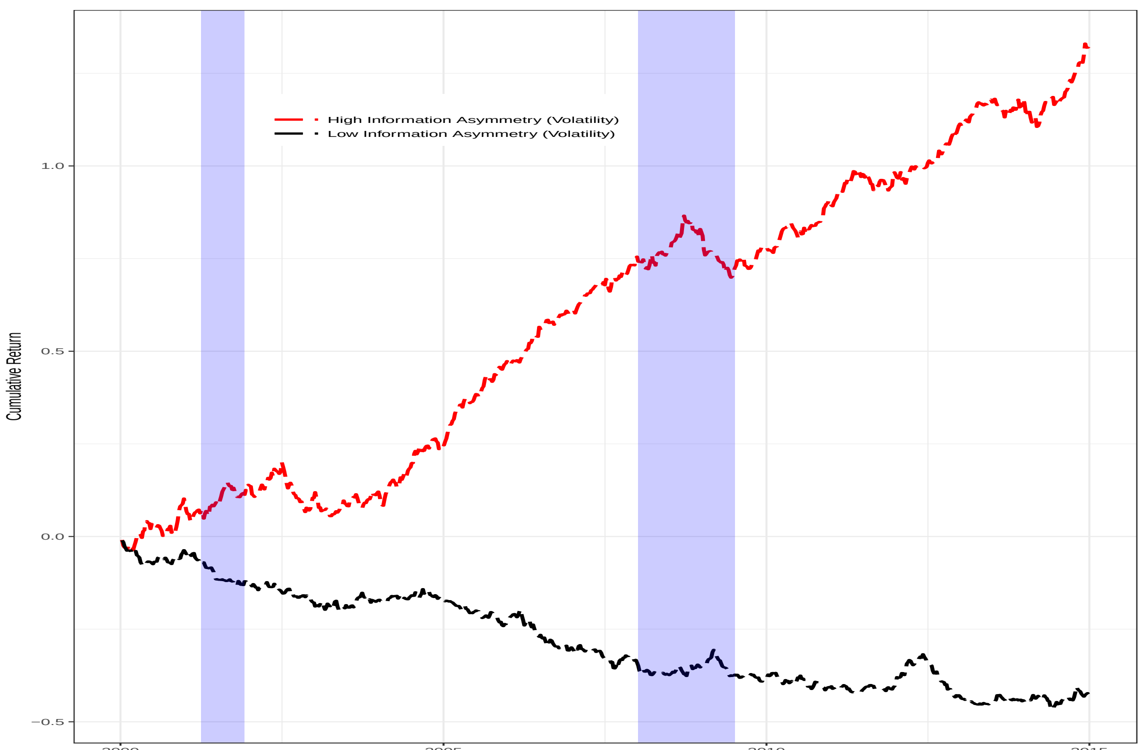


Table 1: **Summary Statistics of Main Variables**

This table provides descriptive statistics for main variables used in the paper. All variables are defined in Appendix. Variables in Panel A are related to earnings announcement with all variables calculated corresponding to the actual earnings announcement date. Variables in Panel B shows weekly industry returns, firm characteristics and Cross Industry News Signal (CIS).

Panel A: 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 1 (continued)

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 2: **Predicting Earnings Using News score**

This table presents results from pooled least square regression of annual earnings (SUE) on cross industry news tone, firm fundamentals and other control variables. The regression model takes the following form:

$$\text{SUE}_{it} = \alpha_i + \sum_{j=1}^{N} \beta_j \text{Industry News}_{j,t-90,t-3} + \gamma^{`}\text{X} + \epsilon_{jt}, \text{where } I_k \neq I_j$$

where the dependent variable, SUE, measures each firm's standardized unexpected earnings following Bernard (1989). Industry $\text{News}_{j,t-90,t-3}$ stands for the news information of cross industries within the period (t-90, t-3) relative to the earnings announcement day. X denotes other explanatory variables which are defined in section 2. In panel A, we only include one target industry news tones as variable of interest, in panel B, we include all industry news tone in the regression and in panel C, we employ adaptive lasso to select most informative predictors. For each panel, we conduct 3 empirical tests by including different control variables. In univarite test, we only include industry news tone as unique predictor, in Tetlock (2008) setting, we calculate all variables used in Tetlock (2008) as control variables as proxy for priced information and market expectation. In All Controls, we include other firm characteristics and public information as additional controls. Coefficients, T-value/P-value and R^2 are presented for each empirical test. We compute clustered standard errors at firm level and use bootstrap method to compute P-value of adaptive lasso.

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 2 (continued)

Panel B: All Industries	SUE					
Empirical Design:	Univariate		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 2 (continued)

Panel C: Adaptive Lasso	SUE					
Empirical Design:	Univariate		Tetlock 2008		All Controls	
	Coef	$P-value$	Coef	$P-value$	Coef	$P-value$
Food	**0.09**	< 0.01			0.01	0.28
Beer	**-0.06**	< 0.01				
Smoke	**-0.03**	0.02				
Games	**-0.03**	0.02				
Books	**-0.03**	0.01				
Hshld	-0.02	0.13				
Clths	0.00	0.31				
Hlth	0.02	0.20				
Chems						
Txtls	**-0.07**	< 0.01				
Cnstr						
Steel						
FabPr	-0.01	0.36				
ElcEq	**-0.07**	< 0.01	**-0.01**	< 0.01	**-0.02**	< 0.01
Autos	**-0.11**	< 0.01	**-0.11**	< 0.01	**-0.08**	< 0.01
Carry						
Mines	**-0.03**	0.01				
Coal	**0.06**	< 0.01				
Oil	**0.12**	< 0.01	**0.01**	< 0.01	**0.02**	< 0.01
Util	**0.11**	< 0.01	**0.04**	< 0.01	**0.05**	< 0.01
Telcm	0.02	0.23				
Servs	**0.04**	0.03				
BusEq	**0.04**	0.02				
Paper	**0.02**	0.10				
Trans	**-0.07**	< 0.01	**-0.02**	< 0.01		
Whlsl	**-0.03**	0.07				
Rtail	**-0.05**	< 0.01				
Meals	**0.05**	< 0.01				
Fin	**0.04**	0.04				
Other	-0.01	0.16				
Year effect	Yes		Yes		Yes	
Firm effect	Yes		Yes		Yes	
N	32,917		32,917		28,206	
adj. R^2 (%)	2.94		20.28		20.85	

Table 3: **Fama-MacBeth regressions of stock returns on CIS**

This table reports the Fama-MacBeth results of regressing stock excess returns on lagged Cross-industry-news signals(CIS), as well as other firm characteristics. CIS is firms' out-of-sample forecasted return based on cross-industry-news tones only. Peer News is average news tone of peer firms within the same industry. Firm News stands for firm specific news tone. Only none-positive CIS observations are included in the regression model. The sample period is between Jan 2000 and 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 4: **Risk-Adjusted Cross-News-Based Trading Strategy Returns**

This table shows the weekly risk-adjusted returns (Alpha) from a CIS trading strategy for three different time periods (2000 to 2014,2000 to 2008 and 2009 to 2014). In each sample period, we first use the market risk to adjust the trading strategy returns for the impact of contemporaneous market. The middle three regressions use Fama-French (1993) three-factor model, size (SMB), and book-to-market (HML) to remove firm characteristic related risk effect. The last three regressions use the Carhart (1997) four-factor model to account for incremental impact of the momentum factor (UMD). Table 4 reports the alpha and loadings from the time-series regression of the long-short CIS portfolio returns on each factors. We assemble the portfolio for the trading strategy at the close of each trading week. We form two equal-weighted portfolios based on the content of each firm's CIS during the prior trading week. We only use stocks with nonpositive CIS to sort portfolio and label all stock with least (most) negative CSI in the previous week's top (bottom) quartile as long (short) leg. We hold both the long and short portfolios for one full trading week and rebalance at the end of the next trading week. The robust t-statistics are in parentheses. ***, ** 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 5: **Sensitivity of News Based Trading Returns to Trading Cost Assumptions**

This table shows estimates of the impact of transaction costs on the cross-industry news and firm specific news based trading strategy's profitability (see the text or Table 4 for cross-industry news strategy details). We recalculate the trading strategy returns for 10 alternative assumptions about a trader's round-trip transaction costs: 1, 2, 3 ... or 10 basis points (bps) per round-trip trade. The abnormal and raw annualized cumulative news based strategy returns for each assumption appear below. The risk-adjustment is based on the full-sample Fama-French three-factor loadings of the news-based portfolio.

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 6: **Sensitivity of News Based Trading Returns to Forecast Horizons**

This table shows estimates of the impact of forecast horizon on the cross-industry news and firm specific news based trading strategy's profitability (see the text or Table 4 for cross-industry news strategy details). We recalculate the trading strategy returns for 9 alternative assumptions about forecast horizon: 2, 3 ... or 10 weeks after CIS. The abnormal and raw annualized cumulative news based strategy returns for each assumption appear below. The risk-adjustment is based on the full-sample Fama-French three-factor loadings of the news-based 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 7: **Alternative Explanation on Cross-News-Based Trading Strategy Returns**

This table shows weekly risk-adjusted returns (Alpha) of cross-industry-news-based trading strategy by considering alternative channels with three different time periods (2000 to 2014, 2000 to 2008, and 2009 to 2014). In each sample period, we first use the Carhart (1997) four-factor model and firm specific news tone based portfolio return to adjust the trading strategy returns of overlapping information with firm specific news. The middle three regressions add on the peer industry news based portfolio return to account for incremental impact of the peer industry news. The last three regressions further control lagged cross-industry-return based portfolio return to account for incremental impact of the priced industry information. Table 7 reports the alpha and loadings from the time-series regression of the long-short cross-industry-news-based portfolio returns on each factors. We assemble the portfolio for the trading strategy at the close of each trading week. We form two equal-weighted portfolios based on CIS during the prior trading week. We only use stocks with nonpositive CIS to sort portfolio and label all stocks with least (most) negative CIS in the previous week's top (bottom) decile as long (short) leg. We hold both the long and short portfolios for one full trading week and rebalance it at the end of the next trading week. The robust t-statistics are in parentheses. ***, ** 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 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 8: **Cross Industry News Portfolio Over High and Low Investor Sentiment Periods**

This table reports average and risk-adjusted returns of cross-industry-news based portfolio return over high and low sentiment periods. We find 3 proxies for sentiment, including BW sentiment index, PLS sentiment index and overall market news sentiment index. We define market news sentiment index as average news tone across industries. A high sentiment week is one if the sentiment index in the previous week is above the median value of the sample period and a low sentiment week otherwise. 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: BW Sentiment					
Low	0.05**	0.06*	0.06**	0.06**	0.04*
	2.09	1.91	2.06	2.09	1.92
High	0.16***	0.16***	0.16***	0.14***	0.11**
	4.82	3.84	3.90	3.61	2.82
Panel B: PLS Sentiment					
Low	0.09***	0.10***	0.10***	0.10***	0.07**
	3.17	2.82	3.04	3.00	2.49
High	0.13***	0.13***	0.13***	0.13***	0.09**
	3.97	3.24	3.31	3.55	2.55
Panel C: Market News					
Low	0.14***	0.14***	0.14***	0.14***	0.12***
	4.14	3.39	3.58	3.93	3.37
High	0.08***	0.09***	0.09***	0.09***	0.05*
	2.96	2.95	3.05	2.94	1.77

Table 9: **Cross Industry News Portfolio Over High and Low Market Uncertainty**

This table reports average and risk-adjusted returns of cross-industry-news based portfolio return over high and low market uncertainty periods. We find 3 proxies for market environment, including VIX, Economic Policy Uncertainty (EPU) and news dispersion index. We define news dispersion index as standard deviation of news tone across individual firms. A high uncertainty week is one if the uncertainty index in the previous week is above the median value of the sample period and a low uncertainty week otherwise. 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 10: **Media News and Analysts' Forecast Behavior**

This table presents results from panel regression of forecast revision or forecast improvement on Cross News Tone, analyst characteristics, firm fundamentals and other control variables. We define Cross News Tone as average of Cross Industry News Tone within analysts' forecast revision periods. The regression model takes the following way:

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

where Y_{ijt} stands for analyst forecast revision or forecast improvement. Forecast revision is the absolute change of analyst forecasts scaled by stock price in the end of last year. Forecast improvement is the current forecast accuracy minus previous forecast accuracy for the same earnings forecast period. We define forecast accuracy as the minus absolute value of difference between actual earnings and analyst forecast. X denotes other explanatory variables which are defined in Section 2. T-statistics are shown in parentheses. Standard errors are clustered on the firm level and robust to heteroskedasticity. ***, ** and * indicate significance at 1%, 5%, and 10%, levels, respectively.

Table 10 (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)
$\text{AR}_{t-252,t-31}$			-0.01633**			-0.01799
			(-2.18)			(-0.67)
$\text{AR}_{t-30,t-3}$			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 11: **Cross Industry News and Institution Fund Flow**

This table presents results from time series regression of weekly industry fund flow on cross-industry-news tone. The regression model takes the following form:

$$\text{Industry Fund Flow}_{jt} \quad = \quad \alpha + \sum_{i=1}^{i=30} \beta_i \text{Industry News Tone}_{\mathrm{i,t-1}} + \sum_{k=1}^{k=30} \beta_k \text{Industry Fund Flow}_{\mathrm{i,t-1}} + \epsilon_{jt},$$

where Industry Fund Flow stands for weekly active institutional fund flow for different industries according to GIC. We map this industry classification to Fama French 30 industries and take cross industry news tone as variable of interst. The first 3 columns in Table 11 count the number of significant industry vraiables that is not overlapped with industry classification of Fund Flow. Column 4 to 7 reports adjusted R^2 that includes lagged industry fund flow alone, lagged industry news tone alone, industry fund flow and industry news tone, all cross industry news tones respectively. Last column reports F statistics using all cross industry news tones

	$p < 0.01$	$p < 0.05$	$p < 0.10$	$R^2_{\text{Fund Flow}}$	$R^2_{\text{Industry News}}$	$R^2_{\text{Fund Flow + Industry News}}$	$R^2_{\text{Cross Industry News}}$	F-test
Fama French 30 Industry: Fin								
RealEstate	1.00	0.00	0.00	4.36	0.07	4.28	0.98	1.17
Insurance	2.00	5.00	2.00	0.64	0.71	1.21	6.01	1.83
DiversifiedFinancials	2.00	6.00	0.00	0.23	0.16	0.34	5.90	2.12
Banks	3.00	4.00	2.00	3.80	2.73	5.46	14.41	3.19
Fama French 30 Industry: Rtail								
Software&Service	2.00	2.00	4.00	-0.01	-0.19	-0.17	6.42	1.89
Rtailing	1.00	2.00	5.00	-0.05	-0.15	-0.16	5.46	1.75
Fama French 30 Industry: Util								
Utilities	1.00	3.00	2.00	0.00	0.18	0.27	3.04	1.57
Fama French 30 Industry: Trans								
Transportation	2.00	4.00	3.00	0.74	-0.22	0.52	7.75	2.09
Fama French 30 Industry: Telcm								
TelecomServices	4.00	4.00	3.00	-0.03	0.42	0.42	14.16	4.20
Fama French 30 Industry: Hshld								
Household&PersonalProducts	1.00	4.00	3.00	-0.04	0.51	0.28	6.14	1.85
Fama French 30 Industry: Hlth								
HealthCareEquip&Services	2.00	0.00	4.00	-0.10	0.50	0.30	5.01	1.69
Fama French 30 Industry: Whlsl								
ConsumerDurable&Apparel	1.00	4.00	4.00	0.81	-0.09	0.86	6.40	1.89
Fama French 30 Industry: Servs								
CommercialServices&Supplies	1.00	5.00	3.00	0.28	-0.16	0.19	5.13	1.70
Fama French 30 Industry: Autos								
Automobiles&Components	3.00	3.00	3.00	6.13	1.06	6.58	10.84	2.58
Fama French 30 Industry: Other								
TechnologyHardware	1.00	3.00	4.00	0.17	-0.24	-0.24	6.18	1.86
Semiconductors	2.00	4.00	2.00	3.77	-0.16	2.97	13.98	3.11
Pharmaceut&Biotechnology	2.00	1.00	5.00	-0.16	-0.20	-0.44	4.69	1.64
Media	2.00	5.00	2.00	0.19	-0.22	-0.24	6.18	1.86
ConsumerServices	2.00	3.00	4.00	0.09	-0.25	-0.36	6.51	1.90
Fama French 30 Industry: Mixed								
Materials	3.00	4.00	2.00	3.14			9.05	2.98
FoodBeverage&Tobacco	2.00	6.00	1.00	2.01			8.92	2.27
Food&DrugRetailing	3.00	5.00	2.00	1.72			9.80	2.41
Energy	2.00	7.00	1.00	7.26			13.16	3.55
Capitalgoods	2.00	4.00	4.00	0.70			7.53	2.06

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

This table reports the Fama-MacBeth results of regressing firm specific news tone on lagged Cross-industry-news signals(CIS), as well as other firm characteristics. Firm News stands for firm specific news tone. CIS is firms' out-of-sample forecasted return based on cross-industry-news tones. Peer News is average news tone of peer firms within the same industry. The sample period is between Jan 2000 and Dec 2014. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable:		Firm 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