# Media Network Based Investors' Attention: A Powerful Predictor of Market Premium

Li Guo*

*Singapore Management University*

Lin Peng†

*City University of New York*

Yubo Tao‡

*Singapore Management University*

Jun Tu§

*Singapore Management University*

June, 2018

*Lee Kong Chian School of Business, Singapore Management University. 50 Stamford Road, Singapore, 178899. E-mail: liguo.2014@pbs.smu.edu.sg.

†Baruch College, City University of New York. Email: lin peng@baruch.cuny.edu

‡School of Economics, Singapore Management University. 90 Stamford Road, Singapore, 178903. E-mail: yubo.tao.2014@phdecons.smu.edu.sg.

§Corresponding author. Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road, Singapore, 178899. Telephone: (+65) 6828 0764. E-mail: tujun@smu.edu.sg. Jun Tu acknowledges that the study was funded through a research grant from Sim Kee Boon Institute for Financial Economics. The usual disclaimer applies. A previous version of this paper has been circulated under the title, "Media Network and Return Predictability".

# Media Network Based Investors' Attention: A Powerful Predictor of Market Premium

**Abstract**

Studies on stock market equity premium predictability mostly examine information-based predictors, such as the traditional fundamental economic variables (hard information) and the recent news tones (soft information). However, investors' attention is largely ignored in the equity premium forecasting literature despite investor attention is crucial on how information is incorporated into stock prices. In this study, we propose a novel predictor, news network triggered attention (NNTA) index, to proxy the abnormal non-shareholders attention by analysing news co-occurrence phenomenon. We show that the *NNTA* index can forecast the market premium significantly and outperform existing state-of-the-art attention proxies and various news-related or information-based predictors.

*JEL Classification*: G11, G12, G41.

*Keywords*: Investors Attention; Network; Return Predictability; Financial Media; News Tones.

Among numerous studies regarding the stock market return predictability, almost all of them are about information-based predictors, mostly using hard information (e.g., fundamental economic variables in Goyal and Welch (2008)) and recently using soft information (e.g., news tones in Tetlock (2007)). However, without investors' attention, information per se is unable to move stock prices. Given that investors' attention has been documented as one of the most important driving forces of stock returns in recent literature, it is surprising that there is a lack of investigation on the impact of investors' attention on market premium forecasting. In this study, we apply media news network to construct a novel investors' attention based predictor, i.e., *news network triggered attention* (NNTA) index, for forecasting market equity premium.

There are evidences suggesting that attention is a scarce resource. An investor may choose to invest in a limited number of stocks and then only pay attention to the information that may influence those stocks they are holding. However, when a news article mentions multiple stocks including the stocks the investor is holding, those stocks not held yet by the investor but mentioned by the article (labelled as *connected stocks*) will be likely to grab the attention of investors as well[1], namely, the *non-shareholders attention*. Unlike the shareholders who may not be short-sales constrained, the non-shareholders would only be able to react to long signals rather than to short signals due to the restriction of short-sales. As a result, the non-shareholders attention will help incorporate good news into the connected stock prices and push it above its fair value. In brief, the more frequent the stocks are co-mentioned by media news, the more non-shareholders attention is triggered, and the higher probability of overvaluation for connected stocks. By aggregating across all the stocks in the market on monthly basis, we formulate an NNTA index using the adjacency matrix in network analysis to gauge the aggregate non-shareholders attention induced by the news co-occurrence.

Empirically, we show our proposed investor-attention-based predictor, NNTA, can forecast the market premium with a significantly negative coefficient and a 5.97% and a

[1]Investors attention can be drawn to a set of stocks mentioned by news (e.g., Barber and Odean (2007) and Yu (2015)).

5.80% monthly in-sample and out-of-sample $R^2$ respectively. In addition, our findings are statistically as well as economically significant even when we control for alternative attention proxies, news-based predictors and hard information predictors, including economic predictors used in Goyal and Welch (2008), media coverage following Fang and Peress (2009), google search index following Da et al. (2011a), the 52-week high following George and Hwang (2004), analyst coverage and trading volume aggregated from individual S&P500 stocks using value weight, and news tones based on Loughran and McDonald (2011) dictionary (Engelberg, 2008; Gurun and Butler, 2012; Hillert et al., 2014; Solomon et al., 2014; Tetlock et al., 2008). In fact, the NNTA can outperform all the existing predictors for both in-sample and out-of-sample. We also examine the performance of our NNTA in predicting returns during the recession and expansion periods, and find that the NNTA obtains larger and positive $R^2$s in both recession and expansion periods comparing with alternative predictors. Moreover, the NNTA index shows significant return predictability only when investors disagreement is high and the short-sales constraint is tight. This is consistent with the intuition that mispricing are more significant when there are high belief divergence and tight short-sales constraint.

We further verify the news attention channel by predicting cross-sectional portfolios and find more frequent news co-occurrence produces lower returns. Indeed, a long-short portfolio based on NNTA index generates a 0.74% monthly return with a monthly Sharpe ratio of 0.14. Moreover, the conventional risk factors such as CAPM, Fama-French (1993) three factors, and Carhart (1997) four factors are unable to explain the alphas generated by our news network triggered attention index. To identify the fundamental source of NNTA, we check the average correlation of Google searches (or Bloomberg searches) between stock pairs. It shows that the more the connected news, the higher the correlation of Google searches (or Bloomberg searches) between the connected firms will be. This result provides direct evidence to support the investor attention interpretation of NNTA.

Lastly, we study the role of centrality score and size weight in modelling investors attention. In network context, a stock attracts investors attention from its connected stocks and the attracted attention may have different loadings on those stocks. Our results reveal that a stock with a low centrality score (small size) tends to be influenced more by media connection, and this impact will be amplified when the stock connects to a high centrality stock (big-size stock) than to a low centrality stock (small-size stock). In particular, a long-short portfolio based on the number of connected news reveals a significant excess return of 1.40% (1.98%) using low-centrality (small-size) stocks that are connected to those high-centrality (big-size) stocks. For the rest types of stocks, we cannot find such strong results. It suggests that the news network triggered investors attention mainly affects a specific type of stocks rather than all the stocks in the market.

Our paper has shed new light upon a different aspect of investors attention. In Peng and Xiong (2006), they documented that investors tend to process more market information than firm-specific information due to limited attention, which result in return co-movement phenomenon. A follow-up work Peng et al. (2007) shows that under both limited attention and attention shifts assumptions, one can explain time-varying asset co-movement. In terms of media attention, Odean (1999) and Barber and Odean (2007) found that individual investors are more likely to trade the stocks that have grabbed their attentions due to limited attention in searching what to trade, especially for buying stocks. Fang and Peress (2009) and Fang et al. (2014) further examined the cross-sectional return predictability and mutual funds' trading and performances using media coverage as proxy of attention-grabbing events, and they also find evidence that both individual and institutional investors subject to limited attention. Different from those papers, we identify an efficient proxy for non-shareholders attention through media network formation. This clear identification address the fact that non-shareholders' trading behaviour is more subject to short-sale constraint comparing to that of shareholders. Therefore, our proxy is more powerful in predicting the market premium than those proxies without distinguishing the roles of the investors.

We also contribute to the literature that studies financial media's role in return predictability. In the past decades, the literature that investigates the media's role in financial markets mainly examines how do the news tones between the lines predict stock prices. Tetlock (2007) presented that the linguistic tone, especially negative tones, can predict market excess returns. Tetlock et al. (2008) further explored the cross-section predictability of returns by processing firm-specific news. Similarly, Zhang et al. (2016) documented a sector specific reaction based on their distilled sentiment measure. Jegadeesh and Wu (2013) further improved Tetlock (2007) with a term weighting scheme based on OLS and Naïve Bayes, and they also find significant return predictability of news articles. Unlike these literature that focuses on extracting investors sentiment in self news, our indices take into account the connected news, and this news connection is shown to possess valuable information for predicting market premium.

Lastly, we contribute to the literature that applies network analysis in financial studies. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) find that economic links among certain individual firms and industries contribute to cross-firm and cross-industry return predictability. They interpret their results as evidence of gradual information diffusion across economically connected firms, in line with the theoretical model of Hong et al. (2007). Rapach et al. (2015) investigate the predictability of industry returns base on a wide range of industrial interdependencies. Different from above literature, we are the first paper to construct the market-wide media network and provide direct evidences on its market return predictability.

The rest of the paper is organized as follows. In section 1, we review the literature exploring media network in financial markets and make some essential assumptions for subsequent analysis. In section 2, we show how to compose a comprehensive measure of media-network-based attention index. Then, we conduct some empirical tests and present our results in section 3. In section 4, we provide economic explanations to our NNTA. We conclude in section 5.

# 1 Media Connection and Media Network

In this section, we review the literatures that study the impact of the media connections and media networks on financial and economic matters, and introduce several reasonable assumptions for constructing the new predictors of stock returns.

Media connection, by definition, is an inter-relationship that is built via news stories which may through explicit mentions or implicit affections. The explicit mentions, also known as media co-occurrence, is the most natural way of formulating the connectivity of two entities. Özgür et al. (2008) first studied the social network inferred from the co-occurrence network of Reuters news. They show that the network exhibits small-world features with power law degree distribution and it provides a better prediction of the ranking on "importance" of people involved in the news comparing to other algorithms. Scherbina and Schlusche (2015) studied the cross-predictability of stock returns by identifying the economic linkage from co-mentions in the news story. They constructed a linkage signal using the weighted average of the connected stock returns and they find that the linked stocks cross-predict one anothers returns in the future significantly, and the predictability increases with the number of the connected news[2].

Apart from the explicit mentions, the connection may also be built through implicit affections. One of the most popular channels is the industrial chain. As shown in Cohen and Frazzini (2008), economic links among certain individual firms and industries contribute significantly to cross-firm and cross-industry return predictability. Rapach et al. (2015) extends the perspective of Cohen and Frazzini (2008) by defining a connection between industries with the predictability of returns. Through these industrial interdependencies, the news that conveys information on one industry will also percolates into the other industries. Further, due to the competitive relation of stocks within the industry, the good (bad) news to one stock will be bad (good) news to its competitors. In addition, business interaction is another important channel that

[2] The *connected news* we are refering to throughout this paper is defined as the news that mentions more than one firm.

transfers news information from one firm to another.

Based on media connections, we can formulate a media network by taking the whole picture of the connected stocks as a undirected graph with news tones or connectivity tagged on each stocks. In network analysis context, all these information can be captured by the *adjacency matrix* or *weighted adjacency matrix*[3]. Apart from adjacency matrix, we also need to make some essential and reasonable assumptions on news arrival and network structures in advance to simplify our analysis.

**Assumption 1** (Random News Arrival)**.** *Connected news arrives randomly and investors have no prior information on the distribution of news arrival.*

In Daley and Green (2012) and Rubin et al. (2017), they presume the news arrival follows some stochastic process or is priori unanticipated. This assumption is reasonable as investors face two tiers of randomness. The first tier of randomness comes from the arrival of firm-specific news event and the second tier comes from the news connections. In reality, a news event is always unpredictable, and even though investors realize a news event will occur, the stocks that the news will mention are still mysterious to the investors.

**Assumption 2** (Multi-degree Network)**.** *The attention that the connected news attracts not only affects the directly connected stocks but also indirectly connected stocks.*

To fit stocks into a network structure based on media connection, the attention attracted by media news could travel through the connected stocks. As a result, attention induced by media connection will not only affect directly linked stocks but also affect those stocks with indirect connections. In this case, the importance of each node (stock) will depend on its connections with all the other nodes (stocks) in this social network. To take this indirect effect into account, we use value weight and eigenvector

[3]In graph theory and computer science, an *adjacency matrix* is a square matrix used to represent an unweighted graph. The elements of the matrix indicate whether pairs of vertices are adjacent or not in the graph. For *weighted adjacency matrix*, it is square matrix used to represent a weighted graph whose edges are tagged with a weight to denote some relationship between the nodes, e.g. distance. The elements of the matrix are just the weight of the edges.

centrality weight to determine the importance of a node in the market. Details will be discussed in the methodology section.

# 2 Data and Methodology

In this section, we will introduce the data sources and explain the methodology for constructing the news network triggered attention index. Then, we introduce the alternative predictors that we can competing with and the corresponding data sources.

## 2.1 News network triggered attention

The data we use for constructing media network is the firm-specific news from the Thomson Reuters News Analytics and Archive dataset ranging from Jan-1996 to Dec-2014. The data contains various types of news, e.g. reviews, stories, analysis and reports etc., about markets, industries and corporations. It also provides three probabilities, namely, $Pos^{NN}$ (the probability of the article being positive), $Neg^{NN}$ (the probability of the article being negative), and $Neu^{NN}$ (the probability of the article being neutral) for all the mentioned firms in each piece of news. These three probabilities sum up to 1 and are being computed from a neural-network-based sentiment engine. In subsequent analysis, we will use $Neg^{NN}$ and $Opt^{NN}$ ($Pos^{NN}$ - $Neg^{NN}$) in addition of soft information predictors.

In this paper, we focus on the firm specific news and define the news that has mentioned at least two stocks as connected news with the others as self-connected news. This dichotomy allows us to construct predictors by aggregating connectivity measures calculated at each node in the network. To construct news-network-based predictors, we start with the *occurrence information matrix*, $\mathcal{M}_t$, that stores the indicators of stocks' occurence in the news in each period:

$$
\mathcal{M}_t = \begin{array}{c} \\ stock_1 \\ stock_2 \\ \vdots \\ stock_N \end{array}
\begin{array}{c}
\begin{array}{cccc} news_1 & news_2 & \cdots & news_{K_t} \end{array} \\
\begin{bmatrix}
Occr_{1,t}^{1} & Occr_{1,t}^{2} & \cdots & Occr_{1,t}^{K_t} \\
Occr_{2,t}^{1} & Occr_{2,t}^{2} & \cdots & Occr_{2,t}^{K_t} \\
\vdots & \vdots & \ddots & \vdots \\
Occr_{N,t}^{1} & Occr_{N,t}^{2} & \cdots & Occr_{N,t}^{K_t}
\end{bmatrix}
\end{array}, \tag{2.1}
$$

where $N$ is the total number of stocks in the sample, $K_t$ is the total number of news of month $t$ which may vary every month, and $Occr_{n,t}^{k}$ equals 1 if stock $n$ is mentioned by news $k$ at time $t$. Based on the occurrence information matrix, we then can obtain the *weighted adjacency matrix*, $\mathcal{W}_t$, that measures the connectivities between the stocks and its strength:

$$
\mathcal{W}_t = \mathcal{M}_t \mathcal{M}_t^{\top} = \begin{array}{c} \\ stock_1 \\ stock_2 \\ \vdots \\ stock_N \end{array}
\begin{array}{c}
\begin{array}{cccc} stock_1 & stock_2 & \cdots & stock_N \end{array} \\
\begin{bmatrix}
w_{11,t} & w_{12,t} & \cdots & w_{1N,t} \\
w_{21,t} & w_{22,t} & \cdots & w_{2N,t} \\
\vdots & \vdots & \ddots & \vdots \\
w_{N1,t} & w_{N2,t} & \cdots & w_{NN,t}
\end{bmatrix}
\end{array}, \tag{2.2}
$$

where $w_{ij,t} = \sum_{k=1}^{K_t} Occr_{i,t}^{k} Occr_{j,t}^{k}$ with $i, j = 1, 2, \cdots, N$. Intuitively, when $i = j$, $w_{ii,t}$ just reflects the media coverage of the stock $i$ at time $t$, and when $i \neq j$, $w_{ij,t}$ accounts for the frequencies of news co-occurrences between stock $i$ and $j$ at time $t$, which can be interpreted to be the connection strength between the stocks or the raw non-shareholders attention score of stock $j$ attracted from stock $i$.

Since the characterization of "non-shareholders" highly depends on the choice of the centering stocks, we need to rescale the raw non-shareholders attention scores by the media coverage of the centering stock so that the non-shareholders attention scores are comparable with each other regardless of the centering stocks. In addition, based on rational expectations, what really moves stock prices is the unexpected non-shareholders attention. In other words, the expectable portion of the non-shareholders attention may be already incorporated in the stock prices. Therefore, assuming that non-shareholders attention is purely random or behave like a random walk, its optimal one-step forecast is the last period's attention level. Then, by taking differences between the realized non-shareholders attention with its forecast, we obtain the *abnormal non-shareholders attention scores* collected in the *adjusted weighted adjacency matrix* as below:

$$\mathcal{AW}_t = \begin{array}{c} \\ stock_1 \\ stock_2 \\ \vdots \\ stock_N \end{array} \begin{array}{c} \begin{array}{cccc} stock_1 & stock_2 & \cdots & stock_N \end{array} \\ \begin{bmatrix} 0 & aw_{12,t} & \cdots & aw_{1N,t} \\ aw_{21,t} & 0 & \cdots & aw_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ aw_{N1,t} & aw_{N2,t} & \cdots & 0 \end{bmatrix} \end{array}, \tag{2.3}$$

where $aw_{ij,t} = w^*_{ij,t} - w^*_{ij,t-1}$ and $w^*_{ij,t} = w_{ij,t}/w_{ii,t}$.

In above formulations, we implicitly assume that each stocks in the news network are equally important such that each stock's abnormal non-shareholders attention takes an equal weight. However, in reality, the more important firms are more easily to seize investors attention. Therefore, we propose to adjust the formulation of occurrence matrix ($\mathcal{M}_t$) with the importance of stocks. In this paper, we adopt two measures to proxy the importance of the stocks. One is the *firm size* and the other one is *eigenvector centrality* which is borrowed from the network analysis.

The centrality is a specialized measure that helps rank the importance of the vertices in the network using the edge information. As introduced in Newman (2010), there are various types of centrality measures applying in network analysis (such as, degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, etc.), and we choose to use eigenvector centrality in our study. Specifically, we first define the adjacency matrix $\mathcal{A}_t$ based on the occurrence information matrix, that is

$$\mathcal{A}_t = \begin{array}{c} \\ stock_1 \\ stock_2 \\ \vdots \\ stock_N \end{array} \begin{array}{c} \begin{array}{cccc} stock_1 & stock_2 & \cdots & stock_N \end{array} \\ \begin{bmatrix} a_{11,t} & a_{12,t} & \cdots & a_{1N,t} \\ a_{21,t} & a_{22,t} & \cdots & a_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1,t} & a_{N2,t} & \cdots & a_{NN,t} \end{bmatrix} \end{array}. \tag{2.4}$$

where $a_{ij,t} = 1$ if $\sum_{k=1}^{K_t} Occr_{i,t}^k Occr_{j,t}^k \neq 0$, and 0 otherwise. Then, we calculate the eigenvector corresponding to the largest eigenvalue ($\lambda_{\max}$) of the adjacency matrix, $\mathbf{x}_t$, which is defined as our centrality score, i.e.,

$$\mathcal{A}_t \mathbf{x}_t = \lambda_{\max} \mathbf{x}_t, \text{ for each } t = 1, 2, \cdots, T, \tag{2.5}$$

where $\mathbf{x}_t = (Ctry_{1,t}, Ctry_{2,t}, \cdots, Ctry_{N,t})'$ and $Ctry_{i,t}$ stands for the eigencvector centrality score of stock $i$ at time $t$.

Unlike the degree centrality awarding one centrality point for every link a node receives, eigenvector centrality thinks not all vertices are equivalent: some are more relevant than others, and, reasonably, endorsements from important nodes count more. In other words, the eigenvector centrality indicates that a node is important if it connects to other important nodes. To understand the centrality connection score better, we take the simple network structure in Figure 1 as an example. Each vertex in the network represents a firm and the edges indicate the media connections induced by news co-occurrence. The degree centrality suggests that firm 1 and 3, firm 2 and 6, or firm 4 and 5 are equally important since they have the same degrees. However, observing that although firm 2 and 6 both have 2 degrees and both connect to firm 1, firm 6 connects to firm 3 which has more degrees, or in other words, more important than firm 4 which is connected to firm 2. Therefore, we should expect firm 6 to be more important than firm 2 in terms of spreading the news as it has more second degree connections. By similar argument, we should also expect firm 5 to take a more central position than firm 4, and firm 1 is more centered than firm 3. Based on the adjacency matrix, we obtain the eigenvector centrality score as the leading eigenvector, which is $[0.5641, 0.2960, 0.5454, 0.1268, 0.2337, 0.4753]$. Evidently, the eigenvector centrality scores fits the situation better in describing the propagation of news.

[Insert Figure 1 here.]

The discussion above explains why we need two weighting schemes for firm's importance: the size or market capitalization determines the attention base while the centrality determines the radiation area of the attention. Based on this argument, we refine the definition of abnormal non-shareholders attention scores with size or centrality weights, i.e.,

$$aw^s_{ij,t} = \begin{cases} Size_{i,t} \times Size_{j,t} \times aw_{ij,t}, & \text{if } s = sz, \\ Ctry_{i,t} \times Ctry_{j,t} \times aw_{ij,t}, & \text{if } s = ctr. \end{cases} \tag{2.6}$$

We finally aggregate the abnormal non-shareholders attention scores of all the stocks in the market to compose *News Network Triggered Attention* (NNTA) index under different weighting schemes,

$$NNTA^s_t = \sum_{i=1}^{N} \sum_{j=1}^{N} aw^s_{ij,t}, \quad s \in \{sz, ctr\}. \tag{2.7}$$

To combine different aspects of the information provided by news network, we then form a composite news network triggered attention index, NNTA, as the simple average of the two standardized individual abnormal non-shareholders attention measures. Since both measures likely contain information about investors' attention as well as idiosyncratic non-attention noise, the averaged news network triggered attention index thus helps to capture the common investor attention component in connected news and diversify away the idiosyncratic noise. To do that, we standardize both $NNTA^{sz}$ and $NNTA^{ctr}$ and then calculate the monthly composite news network triggered attention index, NNTA as simple average of two single factors:

$$NNTA_t = 0.5NNTA_t^{sz} + 0.5NNTA_t^{ctr}. \tag{2.8}$$

In Figure 2, we plot the composite news network triggered attention index and the other two individual media attention indices. As we can see, overall, size-based index shows a similar pattern as centrality weighted attention index. This is because large stocks also tend to be those stocks with high centrality scores and both index reflect media connection induced investor attention. In the meantime, these two indices are still different especially during the expansion period so it is still benefit to combine these two indices together to remove non-attention noise. In addition, the correlation between $NNTA^{sz}$ and $NNTA^{ctr}$ is 0.41 and the composite news network triggered attention index, $NNTA$ shows correlation 0.72 and 0.78 with $NNTA^{sz}$ and $NNTA^{ctr}$ respectively.

[Insert Figure 2 here.]

## 2.2 Alternative predictors

According to Fang and Peress (2009), media coverage has a significant impact on stock returns as a proxy for investor attention. Therefore, to ensure NNTA's predictive power does not purely come from the media coverage, we then calculate the average number of Thomson Reuters news, average number of Dow Jones news, and average number of Wall Street Journal articles to control for the effect of media coverage. We also take first order difference for these predictors to obtain the abnormal media coverages, labelled as $\Delta TRN$, $\Delta DJI$, and $\Delta WSJ$.

Meanwhile, a related type of literature suggests the use of linguistic methods in order to quantify the tone of relevant textual documents (e.g. Engelberg (2008), Gurun and Butler (2012), Hillert et al. (2014), Solomon et al. (2014), Tetlock et al. (2008)).

The limited attention view then predicts that this information has predictive power for the behavior of cognitively overloaded investors suggested by Jacobs (2015). In this case, we construct soft information predictor using value weight to aggregate individual news tones from S&P500 stocks. In particular, negative news tone for individual stock $i$ is in month $t$ is calculated as $Neg = \frac{\#\ of\ Neg\ Words_{i,t}}{Total\ \#\ of\ Words_{i,t}}$, and the optimistic news tone is $Opt = \frac{\#\ of\ Pos\ Words_{i,t} - \#\ of\ Neg\ Words_{i,t}}{Total\ \#\ of\ Words_{i,t}}$, where positive words and negative words follow Loughran and McDonald (2011) dictionary.

Apart from the media news data, we also construct some alternative attention proxies, including google search index (*Google Search*) following Da et al. (2011a), 52-week highest price indicator ($Prc^{High}$) following George and Hwang (2004), level and change of average number of analysts aggregated from individual S&P500 stocks using value weight (*Analyst* and $\Delta$*Analyst*) and the residual of Analyst coverage regressing on Nasdaq index and firm size (*Analyst_r*) following Hong et al. (2000). In addition, we also use the value-weighted trading volume of each stock (*TrdVol*) and the abnormal trading volume ($\Delta$*TrdVol*) to proxy the aggregate market attention.

We further collect 14 economic predictors that are linked directly to economic fundamentals used in Goyal and Welch (2008) from Amit Goyal's website. Specifically, they are the log dividend-price ratio (D/P), log dividend yield (D/Y), log earnings-price ratio (E/P), log dividend payout ratio (D/E), stock return variance (SVAR), book-to-market ratio (B/M), net equity expansion (NTIS), treasury bill rate (TBL), long-term bond yield (LTY), long-term bond return (LTR), term spread (TMS), default yield spread (DFY), default return spread (DFR) and inflation rates (INFL).

Apart from controlling existing attention proxies and economic predictors, we would also like to control for general synchronicity of firm level fundamentals. It is because the stocks co-mentioned by the news are potentially highly correlated in fundamentals. Therefore, we follow Morck et al. (2000) to construct the *Earnings Co-movement Index* (ECI) for controlling fundamental correlations. To construct the index, we first run the regression

$$ROA_i = a_i + b_i \times ROA_m + \epsilon_i, \tag{2.9}$$

for each firm $i$ in each period. $ROA_i$ is a firms returns on assets, calculated as annual after-tax profit plus depreciation over total assets. $ROA_m$ is the value-weighted average of the return on assets for all firms.

$$Earnings\ Co\text{-}movement\ Index = \frac{\sum_i R_i^2(ROA) \times SST_i(ROA)}{\sum_i SST_i(ROA)}, \tag{2.10}$$

where $R_i^2(ROA)$ and $SST_i(ROA)$ are the $R^2$ and the sum of squared total variations derived from regression (2.9) for firm $i$. A higher ECI indicates that the earnings frequently move together.

Moreover, in order to control for investors' belief divergence, we construct the macro disagreement measure by applying principal component to the same set of macro economic variables in Li (2016). We also collect VIX as a complement to macro disagreement. Besides, we compute the short interest ratio (SIR) to check how short-sales constraint affect the return predictability of NNTA.

[Insert Table 1 here.]

From the summary statistics in Table 1 we can observe that the monthly excess market return has a mean of 0.41% and a standard deviation of 4.49%, implying a monthly Sharpe ratio of 0.09. Moreover, most of economic predictors are highly persistent while the excess market return has little autocorrelation. These summary statistics are generally consistent with the literature.

# 3 Predicting Stock Market Returns with News Co-occurrence

In this section, we provide a number of empirical results. Section 3.1 examines the predictability of news network triggered attention index on the aggregate market. Section 3.2 compares the news network triggered attention index with alternative predictors.

Section 3.3 analyses the out-of-sample predictability, and Section 3.4 assesses the cross-sectional predictability of the news network triggered attention index.

## 3.1 Forecasting the market

Consider the standard predictive regression model,

$$R^m_{t+1} = \alpha + \beta X_t + \epsilon_{t+1}, \tag{3.1}$$

where $R^m_{t+1}$ is the excess market return, i.e., the monthly log return on the S&P500 index in excess of the risk-free rate, and $X_t$ is the NNTA indices and other predictors. For comparison, we also run the same in-sample predictive regression with media coverage indices, alternative attention proxies, news tones and earnings comovement index. Specifically, we test the null hypothesis $\mathcal{H}_0 : \beta = 0$, which means *NNTA* has no predictability for stock returns, against the alternative $\mathcal{H}_1 : \beta \neq 0$. Under the null hypothesis, (3.1) reduces to the constant expected return model, $R^m_{t+1} = \alpha + \epsilon_{t+1}$.

[Insert Table 2 here.]

Table 2 reports the results of in-sample predictive regressions. Panel A to Panel E provide the estimation results for the news network triggered attention index, media coverage index, alternative attention proxies, news tones and earnings comovement index. As shown in Panel A, the composite *NNTA* can predict negative returns significantly with an in-sample $R^2$ of 5.97%, and both individual *NNTA* indices significantly predict negative returns with $R^2$ above 2.8% . While for other predictors, only the trading volumes and Wall Street Journal news coverage exhibit significant in-sample predictability. The last three columns report the overall $R^2$ and $R^2$s in expansion and recession periods recorded by NBER. The results show that *NNTA*s achieve largest in-sample $R^2$s amongst all the attention proxies.

Economically, the OLS coefficient suggests that a one-standard-deviation increase in *NNTA* is associated with an approximate 1.09% decrease in expected excess market return for the next month. On the one hand, recall that the average monthly excess market return during our sample period is 0.41%, thus the slope of -1.09% implies that the expected excess market return based on *NNTA* varies by 2.7 times of the magnitude of its average level, which indicates a strong economic impact. On the other hand, if we annualize the 1.09% decrease in one month by the multiplication of 12, the annualized level of 13.08% is somewhat large. In this case, one may interpret this as the model implied expected change that may not be identical to the reasonable expected change of the investors in the market. Empirically, this level is significantly larger than conventional macroeconomic predictors. For example, a one-standard-deviation increase in the D/P ratio, the CAY and the net payout ratio tends to increase the risk premium by 3.60%, 7.39%, and 10.2% per annum, respectively (see, e.g. Lettau and Ludvigson (2001) and Boudoukh et al. (2007)).

Meanwhile, the $R^2$ of *NNTA* with OLS forecast is 5.97%, which is substantially greater than all alternative attention proxies as well as soft/hard information predictors. This implies that if this level of predictability can be sustained out-of-sample, it will be of substantial economic significance (Kandel and Stambaugh (1996)). Campbell and Thompson (2008) show that, given the large unpredictable component inherent in the monthly market returns, a monthly out-of-sample $R^2$ of 0.5% can generate significant economic value and our findings in section 3.3 are consistent with this argument.

Apart from analysing the predictability over the whole sample period, it is also important to check the predictability during business cycles so that we can gain a better understanding of the fundamental driving forces. Following Rapach et al. (2010), we compute the $R^2$ statistics separately for economic expansions ($R^2_{up}$) and recessions ($R^2_{down}$),

$$R^2_c = 1 - \frac{\sum_{t=1}^{T} 1_{\{t \in \mathbb{T}_c\}} \cdot \epsilon_t^2}{\sum_{t=1}^{T} 1_{\{t \in \mathbb{T}_c\}} \cdot (R_t^m - \bar{R}^m)^2}, \quad c \in \{up, down\}, \tag{3.2}$$

where $1_{\{t \in \mathbb{T}_{up}\}}$ ($1_{\{t \in \mathbb{T}_{up}\}}$) is an indicator that takes a value of one when month $t$ is in an NBER expansion (recession) period, i.e., $\mathbb{T}_{up}$ ($\mathbb{T}_{down}$), and zero otherwise; $\epsilon_t$ is the fitted residual based on the in-sample estimates of the predictive regression model in (3.1); $\bar{R}^m$ is the full-sample mean of $R_t^m$; and $T$ is the number of observations for the full sample. Note that, unlike the full-sample $R^2$ statistic, the $R^2_{up}$ ($R^2_{down}$) have no sign restrictions. Columns 4 and 5 of Table 2 report the $R^2_{up}$ and $R^2_{down}$ statistics. It is shown that *NNTA* gains return predictability over the recessions twice as large than over the expansions. In addition, *NNTA* has significant higher return predictability than all the other predictors over the expansion periods, and it only underperforms abnormal WSJ news coverage over the recessions. This confirms that our media network based attention proxy possesses a stable return predictability in all economic environments.

## 3.2 Comparison with economic predictors

In this subsection, we compare the forecasting power of *NNTA*s with alternative predictors and examine whether its forecasting power is driven by omitted attention proxies, soft information, or economic variables related to business cycle fundamentals. Specifically, we examine whether the forecasting power of *NNTA* remains significant after controlling for news coverage, alternative attention proxies, news tones, and economic predictors. To analyse the marginal forecasting power of *NNTA*, we conduct the following bivariate predictive regressions based on *NNTA*s and other predictors,

$$R^m_{t+1} = \alpha + \beta X_t + \phi Z_t + \epsilon_{t+1}, \tag{3.3}$$

where $X_t$ is one of the *NNTA* indices, and $Z_t$ is one of alternative predictors described in section 2.2, and our main interest is the coefficient $\beta$, and to test $\mathcal{H}_0 : \beta = 0$ against $\mathcal{H}_1 : \beta \neq 0$.

[Insert Table 3 here.]

Table 3 shows that the estimates of $\beta$ in (3.3) are negative and stable in magnitude, which is in line with the results of predictive regression (3.1) reported in Table 2. More importantly, $\beta$ remains statistically significant when augmented by other predictors.

These results illustrate that *NNTA* contains sizeable complementary forecasting information beyond what is contained in the media coverage, alternative attention indices, news tones, and economic predictors. Meanwhile, controlling other predictors does not undermine *NNTA*'s impact ($\beta$ remains almost the same magnitude as reported in Table 2), suggesting that the information content of news network based predictors are not overlapping with existing attention proxies.

## 3.3 Out-of-sample forecasts

Despite the in-sample analysis provides more efficient parameter estimates and thus more precise return forecasts by utilizing all available data, Goyal and Welch (2008), among others, argue that out-of-sample tests seem more relevant for assessing genuine return predictability in real time and avoid the over-fitting issue. In addition, out-of-sample tests are much less affected by finite sample biases such as the Stambaugh bias (Busetti and Marcucci (2013)). Hence, it is essential to investigate the out-of-sample predictive performance of media attention indices.

For out-of-sample forecasts at time $t$, we only use information available up to $t$ to forecast stock returns at $t$+1. Following Goyal and Welch (2008), Kelly and Pruitt (2013), and many others, we run the out-of-sample analysis by estimating the predictive regression model recursively based on our news network triggered attention index,

$$\hat{R}^m_{t+1} = \hat{\alpha}_t + \hat{\beta}_t X_{1:t;t}, \tag{3.4}$$

where $X_{1:t;t}$ is the recursively estimated composite*NNTA* index or individual *NNTA* indices, $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the OLS estimates from regressing $\{R^m_{r+1}\}^{t-1}_{r=1}$ with model (3.1) recursively. For comparison purposes, we also carry out out-of-sample test with media coverage, alternative attention indices, news tones and combined economic predictors. The corresponding results are summarized in Panel B to E of Table 4.

To evaluate the out-of-sample forecasting performance, we apply the widely used Campbell and Thompson (2008) $R^2_{OS}$ statistics based on unconstrained forecast and truncated forecast that imposing non-negative equity premium constraint. The unconstrained $R^2_{OS}$ statistic measures the proportional reduction in mean squared forecast error (MSFE) for the predictive regression forecast relative to the historical average benchmark. Goyal and Welch (2008) show that the historical average is a very stringent out-of-sample benchmark, and individual economic variables typically fail to outperform the historical average. To compute $R^2_{OS}$, let $r$ be a fixed number chosen for the initial sample training, so that the future expected return can be estimated at time $t = r+1, r+2, ..., T$. Then, we compute $s = T - r$ out-of-sample forecasts: $\{\hat{R}^m_{t+1}\}_{t=r}^{T-1}$. More specifically, we use first one third data over 1996:01 to 2002:06 as the initial estimation period so that the forecast evaluation period spans over 2002:07 to 2014:12.

$$\hat{R}^2_{OS} = 1 - \frac{\sum_{t=r}^{T-1}(R^m_{t+1} - \hat{R}^m_{t+1})^2}{\sum_{t=r}^{T-1}(R^m_{t+1} - \bar{R}^m_{t+1})^2}, \tag{3.5}$$

where $\bar{R}^m_{t+1}$ denotes the historical average benchmark corresponding to the constant expected return model ($R^m_{t+1} = \alpha + \epsilon_{t+1}$), i.e.,

$$R^m_{t+1} = \frac{1}{t}\sum_{s=1}^{t} R^m_s. \tag{3.6}$$

By construction, the $R^2_{OS}$ statistic lies in the range $(-\infty, 1]$. If $R^2_{OS} > 0$, it means that the forecast $\hat{R}^m_{t+1}$ outperforms the historical average $R^m_{t+1}$ in terms of MSFE.

The statistical significance of the out-of-sample $R^2$s we report is based on MSFE-adjusted statistic of Clark and West (2007) (CW-test hereafter). It tests the null hypothesis that the historical average MSFE is not greater than the predictive regression forecast MSFE against the one-sided (right-tail) alternative hypothesis that the historical average MSFE is greater than the predictive regression forecast MSFE, corresponding to $\mathcal{H}_0 : R^2_{OS} \leq 0$ against $\mathcal{H}_1 : R^2_{OS} > 0$. Clark and West (2007) show that the test has a standard normal limiting distribution when comparing forecasts from the nested models. Intuitively, under the null hypothesis that the constant expected return model generates the data, the predictive regression model produces a noisier forecast than the historical average benchmark as it estimates slope parameters with zero population values. We thus expect the benchmark models MSFE to be smaller than the predictive regression model's MSFE under the null. The MSFE-adjusted statistic accounts for the negative expected difference between the historical average MSFE and predictive regression MSFE under the null, so that it can reject the null even if the $R^2_{OS}$ statistic is negative.

[Insert Table 4 here.]

Panel A of Table 4 show that *NNTA* index generate positive and significant $R^2_{OS}$ statistics (5.80%) and thus delivers a lower MSFE than the historical average. Thus, it is safe to conclude that *NNTA* has strong out-of-sample predictability for market returns, which confirms our conjectures in previous in-sample results (Table 2). Comparing with *NNTA*, all the other predictors show much weaker out-of-sample predictability for market excess returns as shown in Panel B to E. In general, most of the alternative predictors have negative out-of-sample $R^2$s, and their $CW$-statistics are insignificant. Obviously, our news network triggered attention index is a more powerful predictor for market returns amongst other attention proxies and news-related predictors. In addition, the last two columns of Table 4 show that, the predictability of news network triggered attention index are significantly strong and stable across both expansions and recessions.

In summary, out-of-sample analysis shows that, consistent with our previous in-sample results (Tables 2 and 3), news network triggered attention index is a powerful and reliable predictor for the excess market returns, and consistently outperforms other news-related predictors, alternative attention proxies, and combined economic predictors across different sample periods.

### 3.4 Forecasting cross-sectional portfolio

Based on our theory, *NNTA* should predict negative returns given short-sales constraint. The rationale behind is that news co-occurrence attacts non-shareholders attention to connected stocks and this attention generates asymmetric effect between the good news and bad news. Non-shareholders can simply buy the stock to react to the good news while they are not able to short-sales the stock. In this case, an increased news co-occurrence incorporates more good information than bad information into stock price of connected stocks, hence pushing up the prices of those connected stocks above a fair value.

To test the conjecture above, we test the cross-sectional return predictability by sorting on number of connected news[4]. We form 10 equal-weighted portfolios and label the stocks with media attentions in the top (bottom) decile as high (low) attention group. The rest are grouped as median attention group. All portfolios are rebalanced monthly at the close price of next month. The performance of the sorted portfolios are shown in the first column of Table 5. As expected, the low media attention portfolio gains a significant higher alpha than the high media attention portfolio of 0.74% per month ($t$-statistic = 2.15).

[Insert Table 5 here.]

In addition, in Table 5, we test if the alphas generated by media attention, a portfolio that is long stocks with small number of connected news and shortsells stocks with large number of connected news, can be explain by existing factors. We apply CAPM (Markowitz, 1952), Fama-French three-factor model (Fama and French, 1993) and Carhart four-factor model (Carhart, 1997) to dissect the alphas generated by media attention. The results show that media attention portfolio can deliver a high

[4]Directly sorting on *NNTA* is problematic as *NNTA* is constructed by the aggregated change of media co-occurrence, and the change of a market-wide index is different from the aggregation of change for individual stocks cross-sectionally. Meanwhile, it generates missing values by using change of connections. So cross-sectionally, we can only prove our intuition by studying number of connection and weighted scheme seperately.

alpha under all cases. Specifically, the media attention portfolio has Fama and French (1993) abnormal returns of 0.81% per month ($t$-statistic = 2.52). Further adjusting for Carhart (1997) momentum factor, the media attention portfolio earns abnormal returns of 0.62% per month ($t$-statistic = 2.00). These results indicate that connected news indeed captures a different aspect of market excess returns that cannot be explained by conventional market factors.

# 4 Economic Explanations

In this section, we explore the source of predictability of *NNTA* through different channels. First and foremost, we test if higher news co-occurrence induces more frequent search activities, which is an important proxy for investor attention (Da et al., 2011b). Secondly, we examine the performance of *NNTA* under different environments of belief uncertainty and short-sales constraints. Lastly, we justify the economic meaning of using size and centrality weights for constructing *NNTA* by checking how different stock combinations contribute to abnormal returns.

## 4.1 Google and Bloomberg search attention

As discussed in Da et al. (2011b), the attention proxies based on the media occurrence should always make the assumption that if its name was mentioned in the news media, then investors should have paid attention to it. However, news occurrence does not guarantee attention unless investors actually read it. Therefore, Da et al. (2011b) propose using Google search frequency as a direct measure of investor attention.

Respecting the argument in Da et al. (2011b), we then test if news co-occurrences can induce search activities, in order to show our NNTA indeed reflects investor attentions. Firstly, we sort the connected pair stocks into quintiles based on the frequency of news co-occurrence. Then, in each month, we randomly pick up five pairs in each group and calculate the corresponding Google and Bloomberg search volume correlations. The aggregated results are shown below.[5]

[Insert Figure 4 here.]

As shown in Figure 4, the average correlation of Google search and Bloomberg search increase with the news co-occurrences very significantly. Specifically, the average correlations in group with most news co-occurrences are 9% and 17% for Google search and Bloomberg Search respectively. However, the average correlations for group with fewest news co-occurrences are merely 2% and 3% for Google search and Bloomberg Search respectively. These results together provide strong evidence to support the investor attention interpretation of news co-occurrences.

## 4.2 Belief divergence and short-sales constraint

Miller (1977) asserts that the stock prices in equilibrium will reflect only the optimists view and hence will more likely be overvalued when investors have divergent opinions and short-selling is not allowed. Similarly, Hong and Stein (2007) argue that the two key ingredients for explaining stock overpricing behaviour are disagreement stemmed from heterogeneous belief and short-sales constraint. Therefore, to verify these two assumptions, we check the return prediction performance of *NNTA* over high and low environments of belief divergence and of short-sales constraint tightness.

For belief divergence, we construct macro disagreement measure using the same set of macro variables suggested in Li (2016). Instead of using simple average suggested in Li (2016), we apply principal component analysis to extract the most informative factor. In addition, we also use VIX to proxy the investors' belief divergence in the market. For short-sales constraint, we follow Asquith et al. (2005) and use the short interest ratio to proxy the tightness of the short-sales constraint. The in-sample return predictability results under each environment are summarized in Table 6.

[5] For correlation coefficient series of each group, we put them in the appendix, which is available upon request.

[Insert Table 6 here.]

As shown in Table 6, *NNTA*s only show strong return predictability when investors' beliefs are highly divergent and the short-sales constraint is tight. This result justifies our assumptions for news co-occurrence to generate market over-valuation. Actually, media coverage of multiple stocks, in an environment of high belief divergence and tight short-sales constraint, can lead to correlated over-valuation for these stocks. It then spreads to every corner of the market through the network structure and constitutes a market-wide over-valuation proxy. In addition, it shows that weighting scheme is crucial in capturing the attention spreading effect, and we will make a detailed discussion about it in the next subsection.

## 4.3 Centrality and investors attention

In this section, we try to understand the role of centrality scores in affecting attention effect. In the market, there are four types of stocks, namely, stocks with high centrality scores that connect to low centrality stocks (HL), stocks with high centrality scores that connect to high centrality stocks (HH), stocks with low centrality scores that connect to high centrality stocks (LH), and stocks with low centrality scores that are connected with low central stocks (LL). Under media network, a stock attracts investors' attention from its connected stocks. But importantly, the attracted attention would not equally affect all connected stocks. In particular, a stock with a low centrality score tends to be more affected by this connection and this effect will be amplified when the stock is connected to a high centrality stock than that of connecting to a low centrality stock. To understand this argument better, we conduct long-short portfolio within each type of stocks based on the media attention, proxied by the number of connected news.

To balance the level of connections for both long and short stocks in each type of stocks, we independently sort stocks according to the number of connected news, self centrality score (SCS) and average centrality score of connected stocks (CCS).

Specifically, SCS (CCS) classfies stocks into two groups by cutting at median point while the number of connected news divides stocks into 10 deciles. We then report the portfolio return and risk adjusted alpha of attention based trading strategy for each type of stocks. Specifically, we label the group with number of connected news in the top (bottom) decile as high (low) attention group, and our portfolio strategy is to long the stocks in the low attentin group and sell stocks in the high attention group.[6]

[Insert Table 7 here.]

Under this setting, we are able to identify which type of stocks is more sensitive to media connections, and hence contributes to market-wide over-valuation. Table 7 reports the excess portfolio return (risk adjusted portfolio return) of media connection based trading strategy, formed by using different types of stocks. Indeed, not all stocks suffer co-overvaluation – for those stocks with high centrality scores, they are less sensitive to media connection effect with insignificant excess portfolio returns ($t$-statistics are 0.87, 0.94 and 1.62 for HL, HH and LL stocks respectively). Only stocks with low centrality that connect to high centrality stocks (LH) show strong and significant trading profit. The trading strategy generates 1.40% excess return with a $t$-statistic of 3.09. The results cannot be fully explained by conventional risk factors, including CAPM, Fama-French (1993) three factors and Carhart (1997) four factors. As a result, it provides an intuitive way to understand the significance of our centrality weighting scheme, that is, even though the stock itself may receive little attention, but when it links to a giant through news co-occurrence, it will receive excess attention and end up with an over-valuation.

[6] For some periods, a certain type of stocks may not cover any long (short) stocks, we then replace its long (short) excess return with risk free rate (equivalent to long or short a risk free bond)

## 4.4 Size and investors attention

Similarly, in this section, we study the role of value weight in affecting attention effect. We classify stocks into four types, namely, big stocks that are connected to small stocks (*big-connect-small*), big stocks that are connected to big stocks (*big-connect-big*), small stocks that are connected to small stocks (*small-connect-small*), and small stocks that are connected to big stocks (*small-connect-big*). Again, we conduct long-short portfolio within each type of stocks based on the media attention following the same rule we apply for the centrality weight. Table 8 reports the excess portfolio return (risk adjusted portfolio return) of media connection based trading strategy, formed by using different types of stocks. Consistent with our expectation, small stocks that are connected to big stocks tend to be most affected by attention effect. To some extend, the excess portfolio return of connected news based trading strategy achieves 1.98% per monthly with significant $t$-statistic, 2.09 (while is 0.24, -1.43 and -0.38 for *Big-connect-Small*, *Big-connect-Big* and *Small-connect-Small* stocks respectively. The results are also robust after controlling conventional risk factors, including CAPM, Fama-French (1993) three factors and Carhart (1997) four factors. As a result, we provide the economic meaning to the value weight scheme, that is, small stocks are more likely to be affected by media connection, especially when they are linked to big companies. By drawing market attention through big stocks, small stocks receive investors' asymmetry trading behaviour due to short sale constraints, hence contributing to an overall lower market premium. All in all, network structure shows powerful function in transmitting the investor attentions between the stocks and leads to stock mispricing.

[Insert Table 8 here.]

# 5 Conclusions

Investor attention affects market reactions to new information and has been documented as an important driving force of stock returns. Existing literature have constructed predictors using both hard information and soft information, while investors' attention effect seems to be underexplored. Based on media news network, we propose a novel predictor, news network triggered attention index (NNTA), which proxies abnormal non-shareholders attention with media news co-occurrence. In general, we find *NNTA* consistently provides negative return forecasts for both time-series and cross-sectional portfolios. In a sample of S&P500 stocks from 1996 to 2014, we first document *NNTA* can provide significant in-sample and out-of-sample return predictability. Then, we show the return predictability is robust by controlling for other predictors, such as investor sentiment and economic factors. We also provide evidence that *NNTA* captures investor attention by sorting cross-sectional portfolios on news co-occurrence frequencies and by checking the performance of average correlation of Google search and Bloomberg search frequencies.

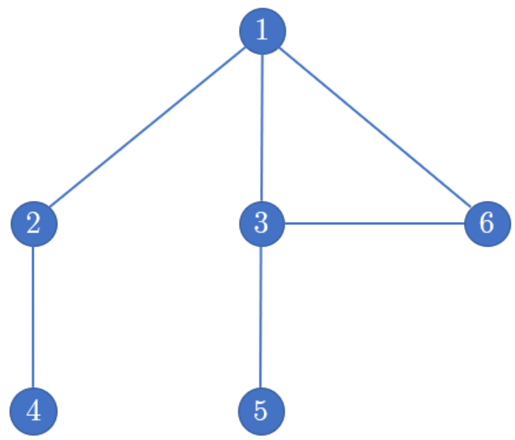


Figure 1: This figure is a simple network example to illustrate how eigenvector centrality differs from degree centrality. Each node in the network represents a company and two nodes are connected when there exists news mentions both of them.

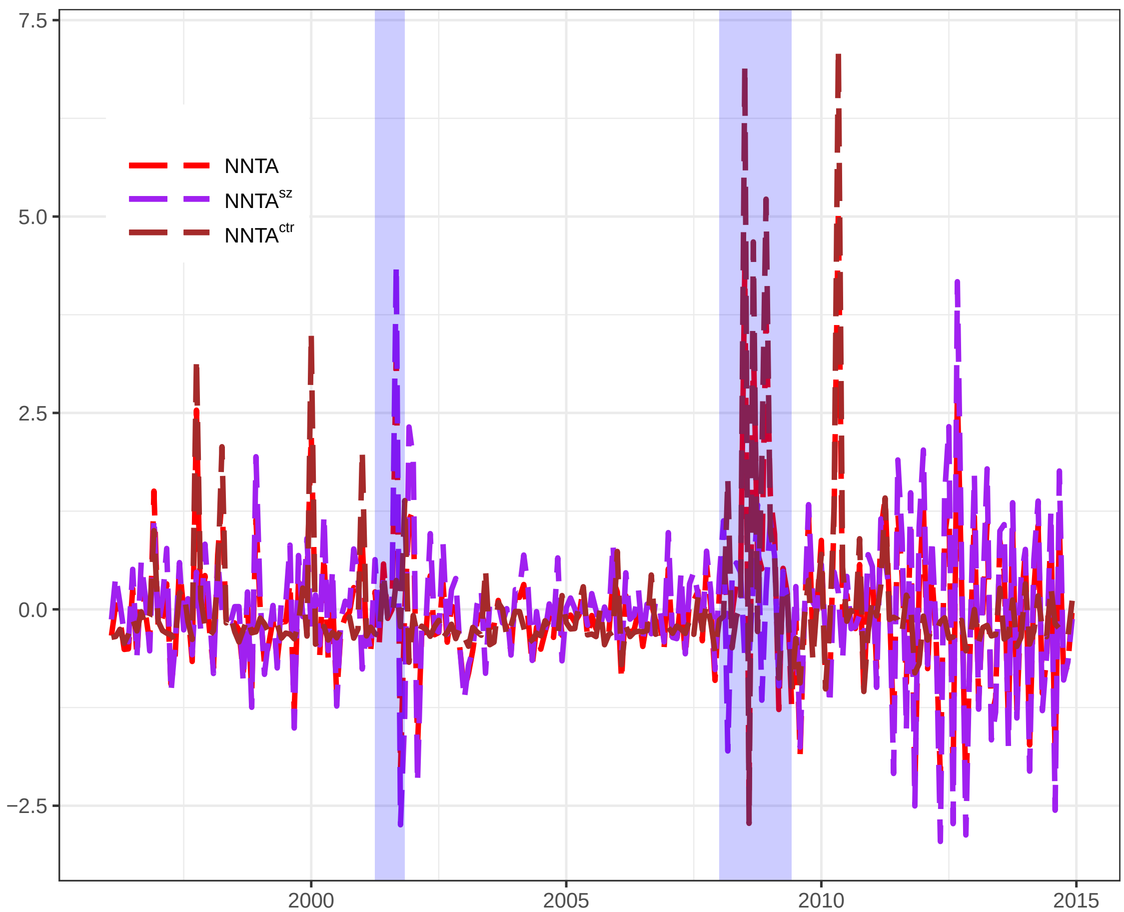


Figure 2: This figure plots the composite news network triggered attention index, size-based news network triggered attention index, and the centrality-based news network triggered attention index. The dashed red line depicts the composite news network triggered attention index, the dashed green line depicts the centrality-based news network triggered attention index, and the dashed purple line depicts the size-based news network triggered attention index. All indices are standardized to have zero mean and unit variance. The shaded periods correspond to NBER-dated recessions. The sample period is 1996:01–2014:12.

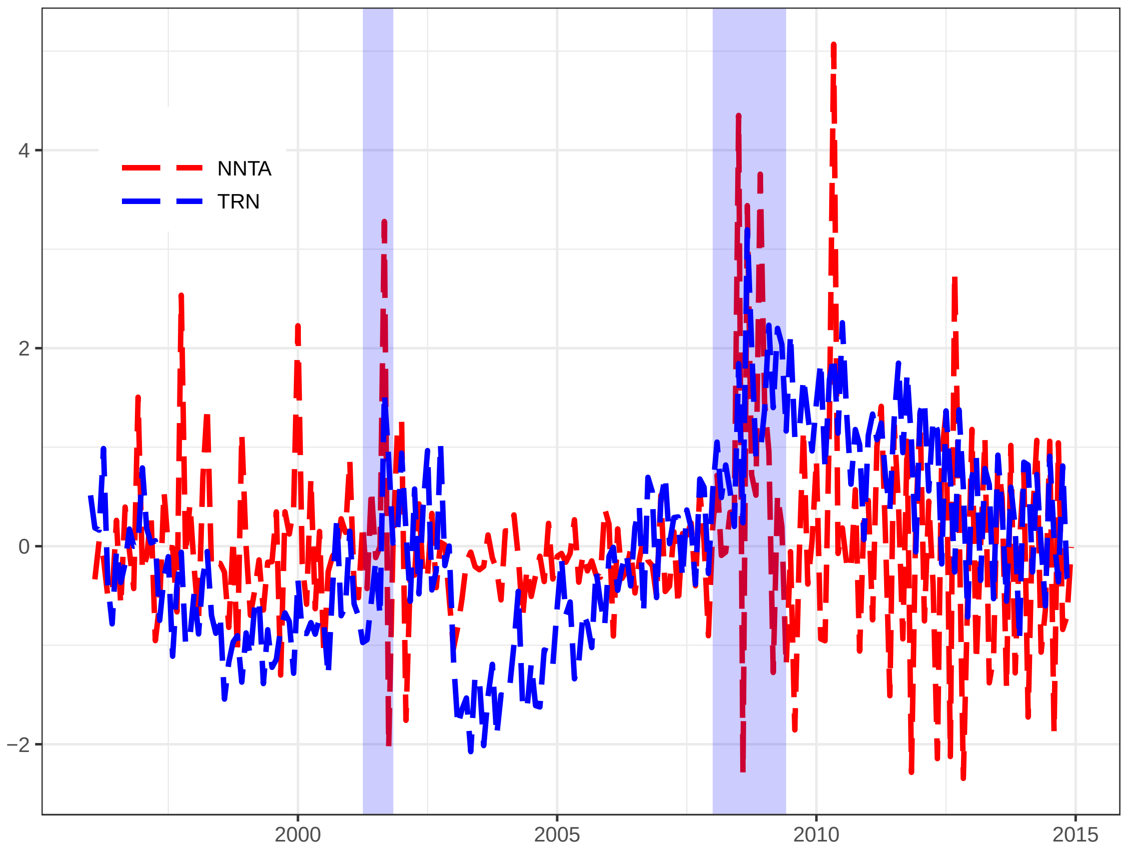


Figure 3: This figure plots the composite news network triggered attention index and value-weighted abnormal number of Thomson Reuters' news. The dashed red line depicts the news network triggered attention index and the dashed blue line depicts the Thomson Reuters News. Both indices are standardized to have zero mean and unit variance. The shaded periods correspond to NBER-dated recessions. The sample period is 1996:01–2014:12.

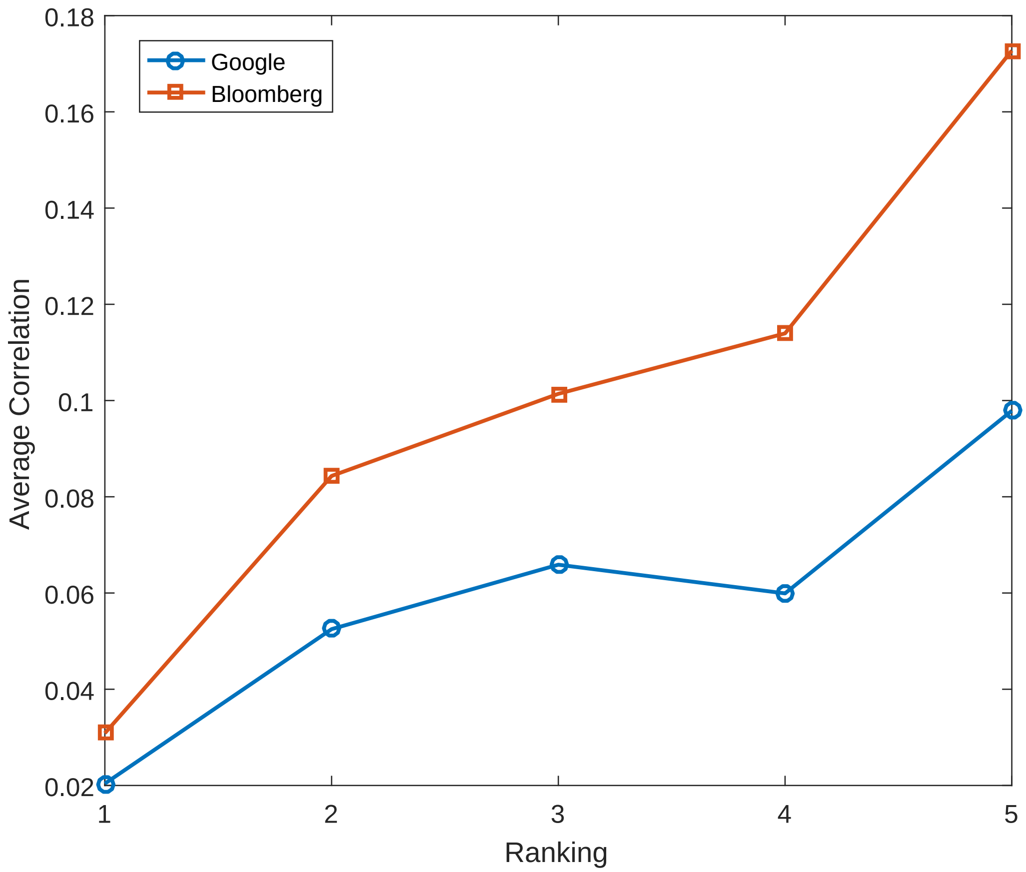


Figure 4: This figure plots the average correlation coefficient of Google and Bloomberg search volumes within each group which is sorted on news attentions. Within each group, the correlation coefficient is calculated monthly using the stock pairs randomly chosen from the 5 sorted groups. The time span is 1996:01–2014:12.

Table 1: Summary Statistics

This table reports summary statistics for the log excess aggregate stock market return defined as the log return on the value-weighted S&P500 stocks in excess of the risk-free rate ($R_m$), risk-free rate ($R_f$), size based news network triggered attention ($NNTA^{sz}$), eigenvector centrality based news network triggered attention, ($NNTA^{ctr}$), and naïvely combined news network triggered attention ($NNTA$); Both level and change of number of Thomson Reuters News/Dow Jones News/Wall Street Journal related to S&P 500 stocks with value weight ($TRN/DJI/WSJ$ and $\Delta TRN/\Delta DJI/\Delta WSJ$)[7]; Google search index (*Google Search*) following Da et al. (2011a), ($Prc^{High}$) following George and Hwang (2004), level and change of average number of analysts aggregated from individual S&P500 stocks using value weight (*Analyst* or $\Delta$*Analyst*), residual of Analyst coverage regressing on size and Nasdaq index following Hong et al. (2000) (*Analyst_r*), value-weighted trading volume (*TrdVol* and $\Delta$*TrdVol*); Negative and optimistic news tones based on Thomson Reuters News Analytics, $Neg^{NN}$ and $Opt^{NN}$ and Loughran and McDonald (2011) dictionary with value weight (*Neg* and *Opt*), Morck et al. (2000) earnings co-movement index (ECI), VIX from CBOE, Asquith et al. (2005) short interest ratio (SIR), and 14 economic variables from Amit Goyals website: the log dividend-price ratio (D/P), the log dividend-yield ratio (D/Y), log earnings-price ratio (E/P), log dividend payout ratio (D/E), stock return variance (SVAR), book-to-market ratio (B/M), net equity expansion (NTIS), Treasury bill rate (TBL), long-term bond yield (LTY) long-term bond return (LTR), term spread (TMS), default yield spread (DFY), default return spread (DFR), inflation rate (INFL). For each variable, the time-series average (Mean), standard deviation (Std. Dev.), skewness (Skew.), kurtosis (Kurt.), minimum (Min.), maximum (Max.), and first-order autocorrelation ($\rho(1)$) are reported. The sample period is 1996:01–2014:12. (Google Search is from 2004:01 – 2014:12)

Table 1 (Continued): Summary Statistics

| *Variable* | Mean | Std. | Skew. | Kurt. | Min. | Max. | $\rho(1)$ |
|---|---|---|---|---|---|---|---|
| **Panel A: Returns** | | | | | | | |
| $R_m$ | 0.0041 | 0.0449 | -0.6565 | 3.9294 | -0.1702 | 0.1077 | 0.0841 |
| $R_f$ | 0.0020 | 0.0018 | 0.2342 | 1.4425 | 0.0000 | 0.0056 | 0.9760 |
| **Panel B: News Network Triggered Attention** | | | | | | | |
| $NNTA$ | 0.001 | 0.730 | 1.347 | 8.254 | -1.700 | 3.676 | -0.180 |
| $NNTA^{sz}$ | 0.000 | 0.002 | 0.258 | 5.929 | -0.005 | 0.007 | -0.357 |
| $NNTA^{ctr}$ | 0.279 | 0.650 | 2.561 | 18.722 | -1.374 | 5.226 | -0.165 |
| **Panel C: Abnormal Media Coverage** | | | | | | | |
| $TRN$ | 3.179 | 0.469 | 0.342 | 2.554 | 2.216 | 4.672 | 0.713 |
| $DJI$ | 22.544 | 17.438 | 0.725 | 2.641 | 0.263 | 71.409 | 0.925 |
| $WSJ$ | 5.553 | 4.421 | 0.616 | 2.185 | 0.136 | 17.087 | 0.938 |
| $\Delta TRN$ | 0.005 | 1.047 | 0.040 | 4.263 | -3.155 | 4.273 | -0.345 |
| $\Delta DJI$ | 0.133 | 6.598 | -0.496 | 11.350 | -36.000 | 29.577 | 0.066 |
| $\Delta WSJ$ | 0.046 | 1.478 | 1.178 | 8.889 | -4.386 | 7.896 | -0.217 |
| **Panel D: Attention Proxies** | | | | | | | |
| *Google Search* | 19.607 | 19.423 | 0.550 | 2.312 | 0.000 | 78.000 | 0.908 |
| $Prc^{High}$ | 0.924 | 0.099 | -1.866 | 6.082 | 0.531 | 0.998 | 0.946 |
| *Analyst* | 25.021 | 1.607 | 0.132 | 1.680 | 22.397 | 27.952 | 0.978 |
| $\Delta Analyst$ | 0.019 | 0.268 | 1.664 | 13.912 | -0.799 | 1.876 | -0.006 |
| *Analyst_r* | -0.169 | 0.040 | -0.184 | 2.710 | -0.266 | -0.060 | 0.953 |
| *TrdVol* | 19.774 | 0.521 | -0.944 | 4.033 | 17.978 | 20.738 | 0.936 |
| $\Delta TrdVol$ | 0.009 | 0.156 | 0.320 | 3.437 | -0.428 | 0.537 | -0.195 |
| **Panel E: Soft Information — News Tones** | | | | | | | |
| *Neg* | 0.006 | 0.002 | 0.562 | 2.900 | 0.003 | 0.010 | 0.726 |
| *Opt* | -0.003 | 0.001 | -0.531 | 3.228 | -0.007 | 0.001 | 0.558 |
| $Neg^{NN}$ | 0.008 | 0.002 | 0.691 | 2.696 | 0.005 | 0.014 | 0.938 |
| $Opt^{NN}$ | 0.005 | 0.002 | 0.978 | 3.568 | 0.001 | 0.011 | 0.888 |
| **Panel F: Hard Information — Economic Predictors** | | | | | | | |
| *ECI* | 0.147 | 0.066 | 0.483 | 2.511 | 0.035 | 0.310 | 0.957 |
| *VIX* | 21.301 | 8.175 | 1.809 | 8.361 | 10.820 | 62.640 | 0.877 |
| *SIR* | 0.015 | 0.003 | 0.429 | 3.196 | 0.010 | 0.022 | 0.956 |
| $D/P$ | -4.016 | 0.399 | 8.664 | 108.696 | -4.524 | 0.953 | 0.305 |
| $D/Y$ | -4.028 | 0.229 | 0.422 | 4.851 | -4.531 | -3.006 | 0.896 |
| $E/P$ | -3.171 | 0.426 | -1.882 | 7.334 | -4.836 | -2.566 | 0.904 |
| $D/E$ | -0.845 | 0.647 | 5.917 | 52.464 | -1.244 | 5.756 | 0.514 |
| *SVAR* | 0.003 | 0.006 | 6.098 | 52.235 | -0.002 | 0.058 | 0.698 |
| $B/M$ | 0.262 | 0.079 | -0.229 | 2.339 | 0.000 | 0.441 | 0.900 |
| *NTIS* | 0.004 | 0.019 | -1.264 | 4.449 | -0.058 | 0.031 | 0.972 |
| *TBL* | 2.435 | 2.130 | 0.200 | 1.390 | 0.010 | 6.170 | 0.985 |
| *LTY* | 4.788 | 1.260 | -0.301 | 2.739 | 0.564 | 7.260 | 0.943 |
| *LTR* | 0.689 | 3.050 | 0.029 | 5.644 | -11.240 | 14.430 | -0.015 |
| *TMS* | 2.354 | 1.406 | -0.453 | 2.710 | -3.226 | 4.530 | 0.903 |
| *DFY* | 0.990 | 0.503 | 0.946 | 17.078 | -2.280 | 3.380 | 0.786 |
| *DFR* | -0.016 | 1.840 | -0.459 | 9.194 | -9.750 | 7.370 | 0.020 |
| *INFL* | 0.002 | 0.004 | 0.534 | 13.781 | -0.019 | 0.029 | 0.325 |

Table 2: Forecasting Market Return with News Network

This table provides in-sample estimation results for the predictive regression of monthly excess market return on news network triggered attention indices, media coverage indices, alternative attention proxies, and news tones

$$R^m_{t+1} = \alpha + \beta X_t + \epsilon_{t+1},$$

where $R^m_{t+1}$ denotes the monthly excess market return (%). *, **, and *** indicate significance at the 10%, 5% and 1% levels respectively. The sample period is 1996:01–2014:12 (Google Search is from 2004:01 – 2014:12).

| *Predictor* | $\hat{\beta}$ | $t$-stat. | $R^2$ | $R^2_{up}$ | $R^2_{down}$ |
|---|---|---|---|---|---|
| **Panel A: News Network Triggered Attention** | | | | | |
| *NNTA* | -1.092*** | -3.762 | 5.969 | 3.831 | 7.045 |
| $NNTA^{sz}$ | -0.750*** | -2.543 | 2.817 | 2.150 | 3.807 |
| $NNTA^{ctr}$ | -0.833*** | -2.834 | 3.475 | 2.105 | 3.179 |
| **Panel B: Abnormal Media Coverage** | | | | | |
| *TRN* | -0.200 | -0.667 | 0.199 | 0.038 | 0.774 |
| *DJI* | 0.264 | 0.882 | 0.348 | 0.221 | 0.262 |
| *WSJ* | 0.153 | 0.511 | 0.117 | 0.313 | 0.335 |
| $\Delta TRN$ | -0.262 | -0.877 | 0.344 | 0.053 | 1.958 |
| $\Delta DJI$ | 0.035 | 0.117 | 0.006 | 0.269 | 4.363 |
| $\Delta WSJ$ | -0.624** | -2.104 | 1.947 | 0.141 | 14.146 |
| **Panel C: Alternative Attention Proxy** | | | | | |
| *Google Search* | -0.260 | -0.868 | 0.337 | 0.006 | 0.047 |
| $Prc^{High}$ | 0.224 | 0.748 | 0.250 | 0.012 | 6.609 |
| *Analyst* | -0.050 | -0.165 | 0.012 | 0.425 | 0.248 |
| $\Delta Analyst$ | -0.124 | -0.415 | 0.077 | 0.003 | 4.550 |
| *Analyst_r* | 0.188 | 0.627 | 0.176 | 0.634 | 0.192 |
| *TrdVol* | -0.587** | -1.977 | 1.722 | 0.732 | 0.607 |
| $\Delta TrdVol$ | -0.666** | -2.246 | 2.212 | 3.474 | 2.462 |
| **Panel D: Soft Information — News Tones** | | | | | |
| *Neg* | -0.245 | -0.810 | 0.296 | 0.273 | 0.006 |
| *Opt* | 0.299 | 0.992 | 0.443 | 0.011 | 0.035 |
| $Neg^{NN}$ | -0.291 | -0.965 | 0.415 | 1.070 | 0.905 |
| $Opt^{NN}$ | 0.456 | 1.524 | 1.031 | 1.171 | 0.039 |
| **Panel E: Hard Information — Earnings Comovement** | | | | | |
| *ECI* | -0.021 | -0.069 | 0.002 | 0.118 | 6.456 |

## Table 3: Comparison with Alternative Predictors

This table provides in-sample estimation results for the bivariate predictive regression of monthly excess market return on one of media coverage, alternative attention proxies, news tones, or 14 economic predictors, $Z_t$, and on the news network triggered attention index, $X_t$.

$$R^m_{t+1} = \alpha + \beta X_t + \phi Z_t + \epsilon_{t+1},$$

where $R^m_{t+1}$ denotes the monthly excess market return (%). The significance of the estimates are based on Newey-West $t$-statistics. *, **, and *** indicate significance at the 10%, 5% and 1% levels respectively. The sample period is 1996:01–2014:12 (Google Search is from 2004:01 – 2014:12).

| | *NNTA* | | | $NNTA^{sz}$ | | | $NNTA^{ctr}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| *Predictor* | $\hat{\beta}$ | $\hat{\phi}$ | $R^2$ | $\hat{\beta}$ | $\hat{\phi}$ | $R^2$ | $\hat{\beta}$ | $\hat{\phi}$ | $R^2$ |
| **Panel A: Abnormal Media Coverage** | | | | | | | | | |
| *TRN* | -1.122*** | 0.102 | 6.019 | -0.740*** | -0.092 | 2.868 | -0.836*** | 0.007 | 3.471 |
| *DJI* | -1.110*** | 0.320 | 6.480 | -0.752*** | 0.265 | 3.175 | -0.859*** | 0.330 | 4.012 |
| *WSJ* | -1.121*** | 0.267 | 6.321 | -0.761*** | 0.189 | 3.003 | -0.859*** | 0.243 | 3.762 |
| $\Delta TRN$ | -1.183*** | 0.219 | 6.169 | -0.763*** | 0.026 | 2.829 | -0.816*** | -0.090 | 3.510 |
| $\Delta DJI$ | -1.113*** | 0.163 | 6.101 | -0.763*** | 0.107 | 2.883 | -0.842*** | 0.096 | 3.517 |
| $\Delta WSJ$ | -1.003*** | -0.391 | 6.691 | -0.646** | -0.481 | 3.922 | -0.773*** | -0.536* | 4.885 |
| **Panel B: Attention Proxies** | | | | | | | | | |
| *Google Search* | -1.082*** | -0.189 | 6.147 | -0.748*** | -0.244 | 3.124 | -0.820*** | -0.197 | 3.664 |
| $Prc^{High}$ | -1.083*** | 0.104 | 6.024 | -0.755*** | 0.230 | 3.090 | -0.820*** | 0.085 | 3.506 |
| *Analyst* | -1.109*** | -0.152 | 6.084 | -0.753*** | -0.052 | 2.840 | -0.855*** | -0.160 | 3.595 |
| $\Delta Analyst$ | -1.113*** | -0.217 | 6.203 | -0.761*** | -0.161 | 2.956 | -0.848*** | -0.183 | 3.638 |
| *Analyst_r* | -1.114*** | 0.266 | 6.323 | -0.751*** | 0.177 | 2.982 | -0.868*** | 0.288 | 3.879 |
| *TrdVol* | -1.021*** | -0.401 | 6.741 | -0.712** | -0.533* | 4.233 | -0.750*** | -0.447 | 4.428 |
| $\Delta TrdVol$ | -0.987*** | -0.261 | 6.253 | -0.590* | -0.454 | 3.719 | -0.718*** | -0.498* | 4.638 |
| **Panel C: Soft Information — News Tones** | | | | | | | | | |
| *Neg* | -1.104*** | -0.263 | 6.340 | -0.760*** | -0.246 | 3.161 | -0.841*** | -0.265 | 3.807 |
| *Opt* | -1.095*** | 0.281 | 6.387 | -0.760*** | 0.300 | 3.306 | -0.829*** | 0.279 | 3.846 |
| $Neg^{NN}$ | -1.108*** | -0.337 | 6.527 | -0.767*** | -0.326 | 3.346 | -0.839*** | -0.305 | 3.927 |
| $Opt^{NN}$ | -1.092*** | 0.451 | 6.973 | -0.760*** | 0.470 | 3.914 | -0.824*** | 0.438 | 4.416 |
| **Panel D: Hard Information — Economic Predictors** | | | | | | | | | |
| *ECI* | -1.095*** | 0.019 | 5.972 | -0.753*** | -0.009 | 2.827 | -0.835*** | 0.011 | 3.472 |
| *D/P* | -1.151*** | 1.202** | 8.131 | -0.748*** | 1.010 | 4.362 | -0.926*** | 1.248** | 5.776 |
| *D/Y* | -1.129*** | 0.728*** | 8.371 | -0.738*** | 0.651 | 4.749 | -0.904*** | 0.760** | 6.070 |
| *E/P* | -1.089*** | 0.195 | 6.146 | -0.754*** | 0.234 | 3.081 | -0.826*** | 0.187 | 3.633 |
| *D/E* | -1.104*** | 0.195 | 6.070 | -0.753*** | 0.084 | 2.845 | -0.851*** | 0.218 | 3.595 |
| *SVAR* | -0.997*** | -0.445 | 6.906 | -0.719*** | -0.626** | 4.763 | -0.709*** | -0.477 | 4.520 |
| *B/M* | -1.109*** | 0.361 | 6.588 | -0.748*** | 0.301 | 3.257 | -0.863*** | 0.382 | 4.160 |
| *NTIS* | -1.042*** | 0.476 | 7.087 | -0.736*** | 0.569* | 4.436 | -0.771*** | 0.491* | 4.651 |
| *TBL* | -1.102*** | -0.229 | 6.231 | -0.751*** | -0.184 | 2.995 | -0.849*** | -0.243 | 3.763 |
| *LTY* | -1.093*** | -0.319 | 6.453 | -0.745*** | -0.305 | 3.267 | -0.841*** | -0.341 | 4.022 |
| *LTR* | -1.093*** | 0.113 | 6.034 | -0.751*** | 0.016 | 2.828 | -0.865*** | 0.237 | 3.746 |
| *TMS* | -1.099*** | 0.079 | 5.999 | -0.753*** | 0.019 | 2.828 | -0.841*** | 0.081 | 3.501 |
| *DFY* | -1.075*** | -0.123 | 6.029 | -0.740*** | -0.295 | 3.177 | -0.807*** | -0.149 | 3.556 |
| *DFR* | -1.094*** | 0.331 | 6.514 | -0.809*** | 0.437 | 3.761 | -0.805*** | 0.228 | 3.724 |
| *INFL* | -1.091*** | 0.156 | 6.068 | -0.760*** | 0.214 | 3.010 | -0.827*** | 0.120 | 3.528 |

## Table 4: Out-of-sample Forecasting

This table reports the out-of-sample performances of various measures of News Network Triggered Attention Indices in predicting the monthly excess market return. Panel A provides the results using the NNTA indices. Panel B are results of abnormal media coverage. Panel C are results using alternative attention proxies. Panel D reports results using news tones and Panel E is the result of combined economic predictors by Rapach et al. (2010) and earning comovement index by Morck et al. (2000). All of the predictors and regression slopes are estimated recursively using the data available at the forecast formation time $t$. $R^2_{OS}$ is the out-of-sample $R^2$ with no constraints. CW-test is the Clark and West (2007) MSFE-adjusted statistic calculated according to prevailing mean model. $R^2_{OS,up}$ ($R^2_{OS,down}$) statistics are calculated over NBER-dated business-cycle expansions (recessions) based on the no constraint model. The out-of-sample evaluation period is 2002:07–2014:12 (Google Search is from 2008:01 – 2014:12).

| *Predictor* | $R^2_{OS}$ | CW-test | $p$-value | $R^2_{OS,up}$ | $R^2_{OS,down}$ |
|---|---|---|---|---|---|
| **Panel A: News Network Triggered Attention Indices** | | | | | |
| *NNTA* | 5.800 | 2.658 | 0.004 | 4.496 | 8.184 |
| $NNTA^{sz}$ | 2.607 | 2.549 | 0.005 | 0.786 | 5.936 |
| $NNTA^{ctr}$ | 2.227 | 1.295 | 0.098 | 3.812 | -0.670 |
| **Panel B: Abnormal Media Coverage** | | | | | |
| *TRN* | -3.667 | -0.344 | 0.635 | -6.413 | 1.351 |
| *DJI* | -0.217 | -0.109 | 0.544 | -0.291 | -0.083 |
| *WSJ* | -5.251 | 0.291 | 0.385 | -7.088 | -1.892 |
| $\Delta TRN$ | -1.863 | -0.392 | 0.653 | -0.514 | -4.328 |
| $\Delta DJI$ | -1.051 | -0.939 | 0.826 | -1.048 | -1.057 |
| $\Delta WSJ$ | -3.001 | 0.279 | 0.390 | -1.863 | -5.081 |
| **Panel C: Attention Proxies** | | | | | |
| *Google Search* | 0.859 | 1.097 | 0.136 | 3.775 | -2.196 |
| $Prc^{High}$ | -2.537 | -0.032 | 0.513 | -1.950 | -3.610 |
| *Analyst* | -2.362 | -0.766 | 0.778 | -1.248 | -4.398 |
| $\Delta Analyst$ | -0.412 | -0.447 | 0.673 | -1.163 | 0.960 |
| *Analyst_r* | -0.888 | 0.235 | 0.407 | -0.353 | -1.865 |
| *TrdVol* | -0.653 | 0.784 | 0.217 | -6.682 | 10.368 |
| $\Delta TrdVol$ | 0.778 | 1.098 | 0.136 | 3.000 | -3.283 |
| **Panel D: Soft Information — News Tones** | | | | | |
| *Neg* | -2.045 | -0.171 | 0.568 | -2.825 | -0.618 |
| *Opt* | -1.102 | 0.002 | 0.499 | -1.571 | -0.246 |
| $Neg^{NN}$ | -0.833 | 0.006 | 0.498 | -1.108 | -0.330 |
| $Opt^{NN}$ | 0.139 | 0.567 | 0.285 | 0.228 | -0.022 |
| **Panel E: Hard Information — Combined Economic Predictors** | | | | | |
| *ECI* | -1.225 | -0.077 | 0.531 | -1.197 | -1.277 |
| *Mean* | -0.669 | 0.003 | 0.499 | -0.330 | 1.350 |
| *Median* | 0.052 | 0.224 | 0.411 | 0.178 | 2.423 |
| *Trimmed Mean* | -0.493 | -0.001 | 0.500 | -0.328 | 1.836 |
| *DMSPE*, $\theta = 1.0$ | -0.693 | 0.020 | 0.492 | -0.211 | 1.130 |
| *DMSPE*, $\theta = 0.9$ | -0.606 | 0.097 | 0.461 | -0.239 | 1.370 |

Table 5: Performance of Sorted Decile Portfolios Based on Media Co-occurrence

This table reports excess portfolio return and risk adjusted alpha of investment strategies based on number of connected news in last month. The sample period is from Jan, 1996 to Dec, 2014. We first sort stocks into 10 deciles according to firms' number of connected news and label all stocks with number of connected news in the top (bottom) decile as short (long) group. We hold each group of stocks for 1 month and rebalance them at the close price of next month. Three types of risk factors are considered: CAPM, Fama-French (1993) three-factor model, including size (SMB), and book-to-market (HML) and Carhart (1997) four-factor model to account for incremental impact of the momentum factor. $t$-statistics are reported below the portfolio return (risk adjusted alpha).

| *Portfolios* | $R_m$ | CAPM | FF-3 | Cahart-4 |
|---|---|---|---|---|
| Long | 0.90% | 0.30% | 0.13% | 0.21% |
| | (2.59) | (1.77) | (0.98) | (1.61) |
| 2 | 0.88% | 0.26% | 0.11% | 0.20% |
| | (2.53) | (1.73) | (0.93) | (1.85) |
| 3 | 1.05% | 0.44% | 0.32% | 0.44% |
| | (3.04) | (2.77) | (2.23) | (3.32) |
| 4 | 1.06% | 0.44% | 0.31% | 0.43% |
| | (2.96) | (2.57) | (2.06) | (3.03) |
| 5 | 0.79% | 0.15% | 0.04% | 0.20% |
| | (2.13) | (0.82) | (0.22) | (1.28) |
| 6 | 0.83% | 0.18% | 0.11% | 0.22% |
| | (2.28) | (1.16) | (0.76) | (1.57) |
| 7 | 0.92% | 0.21% | 0.13% | 0.27% |
| | (2.28) | (1.14) | (0.73) | (1.60) |
| 8 | 0.60% | -0.07% | -0.11% | -0.01% |
| | (1.51) | (-0.31) | (-0.53) | (-0.06) |
| 9 | 0.68% | -0.09% | -0.11% | 0.11% |
| | (1.50) | (-0.37) | (-0.49) | (0.53) |
| Short | 0.16% | -0.67% | -0.67% | -0.41% |
| | (0.32) | (-2.28) | (-2.31) | (-1.54) |
| Long - Short | **0.74%** | **0.96%** | **0.81%** | **0.62%** |
| | (2.15) | (2.92) | (2.52) | (2.00) |

Table 6: Return Predictability under Different Level of Belief Divergence and Short-sales Constraints

This table provides in-sample estimation results for the predictive regression of monthly excess market return on news network triggered attention indices over high and low belief divergence environment as well as high and low short-sales constraint periods. We use macro disagreement and VIX as proxy of belief divergence and use value weighted short interest ratio of S&P500 stocks as proxy for short-sales contraint. A high belief divergence (short-sales constraint) indicator equals one if the belief divergence index (short interest ratio) in the previous month is above the median value of the sample period and 0 otherwise. The sample period is 1996:01–2014:12. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

| *Predictor* | High | | | Low | | |
|---|---|---|---|---|---|---|
| | $\hat{\beta}$ | $t$-stat. | $R^2$ | $\hat{\beta}$ | $t$-stat. | $R^2$ |
| Panel A: Macro Disagreement | | | | | | |
| $NNTA$ | -1.148*** | -3.372 | 0.090 | -0.913 | -1.371 | 0.017 |
| $NNTA^{sz}$ | -0.764** | -2.167 | 0.039 | -0.754 | -1.190 | 0.013 |
| $NNTA^{ctr}$ | -0.914*** | -2.635 | 0.057 | -0.503 | -0.739 | 0.005 |
| Panel B: VIX | | | | | | |
| $NNTA$ | -1.434*** | -3.242 | 0.086 | -0.168 | -0.525 | 0.003 |
| $NNTA^{sz}$ | -1.475*** | -2.792 | 0.065 | -0.079 | -0.318 | 0.001 |
| $NNTA^{ctr}$ | -0.747* | -1.919 | 0.032 | -0.710 | -0.883 | 0.007 |
| Panel C: Short Interest Ratio | | | | | | |
| $NNTA$ | -1.194*** | -3.595 | 0.103 | -0.929 | -1.607 | 0.023 |
| $NNTA^{sz}$ | -0.729** | -1.983 | 0.034 | -0.808 | -1.625 | 0.024 |
| $NNTA^{ctr}$ | -1.004*** | -3.106 | 0.079 | -0.348 | -0.512 | 0.002 |

Table 7: Risk Adjusted Alphas of Attention-based Trading Strategies under Centrality Weights

We independently sort stocks according to the number of connected news, self centrality score (SCS) and average centrality score of connected stocks (CCS). SCS (CCS) classfies stocks into 2 groups by cutting at median point while the number of connected news divides stocks into 5 groups. We then report the portfolio return and risk adjusted alpha of attention based trading strategy under 4 types of stocks, including stocks with high centrality scores that connect to low centrality stocks (*high-connect-low*), stocks with high centrality scores that connect to high centrality stocks (*high-connect-high*), stocks with low centrality scores that connect to low centrality stocks (*low-connect-low*), and stocks with low centrality scores that connect to high centralilty stocks (*low-connect-high*). The trading strategy labels all stocks with number of connected news in the top (bottom) group as high (low) attention group and the portfolio is formed by buying stocks in the low attentin group while selling stocks in the high attention group in last month. For some periods, when a certain type of stocks do not meet any long (short) stocks, we replace the long (short) excess return with risk free rate. We then hold this portfolio for 1 month and rebalance stocks at the close price of next month. Three types of risk factors are considered to find risk adjusted alpha: CAPM, Fama-French (1993) three-factor model, including size (SMB), and book-to-market (HML) and Carhart (1997) four-factor model to account for incremental impact of the momentum factor. $t$-statistics are reported in parentheses below the portfolio return (risk adjusted alpha). The sample period is 1996:01–2014:12

| *Portfolios* | $R_m$ | CAPM | FF-3 | Cahart-4 |
|---|---|---|---|---|
| *High-connect-Low* | 0.75% | 0.78% | 0.54% | 0.26% |
| | (0.87) | (0.89) | (0.61) | (0.29) |
| *High-connect-High* | 0.26% | 0.22% | 0.24% | 0.24% |
| | (0.94) | (0.77) | (0.83) | (0.83) |
| *Low-connect-Low* | 1.28% | 1.04% | 0.99% | 1.14% |
| | (1.62) | (1.31) | (1.24) | (1.42) |
| *Low-connect-High* | **1.40%** | **0.75%** | **0.57%** | **0.61%** |
| | (3.09) | (2.38) | (1.89) | (2.02) |

Table 8: Risk Adjusted Alphas of Attention-based Trading Strategies under Value Weights

We independently sort stocks according to the number of connected news, firm self value weight (SVW) and average value weight of connected stocks (CVW). SVW (CVW) classfies stocks into 2 groups by cutting at median point while the number of connected news divides stocks into 5 groups. We then report the portfolio return and risk adjusted alpha of attention based trading strategy under 4 types of stocks, including big stocks that are connected to small stocks (*big-connect-small*), big stocks that are connected to big stocks (*big-connect-big*), small stocks that are connected to small stocks (*small-connect-small*), and small stocks that are connected to big stocks (*small-connect-big*). The trading strategy labels all stocks with number of connected news in the top (bottom) group as high (low) attention group and the portfolio is formed by buying stocks in the low attentin group while selling stocks in the high attention group in last month. For some periods, when a certain type of stocks do not meet any long (short) stocks, we replace the long (short) excess return with risk free rate. We then hold this portfolio for 1 month and rebalance stocks at the close price of next month. Three types of risk factors are considerred to find risk adjusted alpha: CAPM, Fama-French (1993) three-factor model, including size (SMB), and book-to-market (HML) and Carhart (1997) four-factor model to account for incremental impact of the momentum factor. $t$-statistics are reported in parentheses below the portfolio return (risk adjusted alpha). The sample period is 1996:01–2014:12

| *Portfolios* | $R_m$ | CAPM | FF-3 | Cahart-4 |
|---|---|---|---|---|
| *Big-connect-Small* | 0.08% | 0.23% | 0.17% | -0.05% |
| | (0.24) | (0.67) | (0.51) | (-0.14) |
| *Big-connect-Big* | -0.83% | -0.34% | -0.39% | -0.58% |
| | (-1.43) | (-0.65) | (-0.72) | (-1.08) |
| *Small-connect-Small* | -0.27% | -0.29% | -0.19% | -0.22% |
| | (-0.38) | (-0.39) | (-0.25) | (-0.29 ) |
| *Small-connect-Big* | **1.98%** | **2.10%** | **2.17%** | **1.85%** |
| | (2.09) | (2.20) | (2.26) | (1.93) |

# References


Asquith, P., P. A. Pathak, and J. R. Ritter (2005): "Short interest, institutional ownership, and stock returns," *Journal of Financial Economics*, 78, 243–276.

Baker, M. and J. Wurgler (2006): "Investor sentiment and the cross-section of stock returns," *The Journal of Finance*, 61, 1645–1680.

Barber, B. M. and T. Odean (2007): "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors," *The Review of Financial Studies*, 21, 785–818.

Boudoukh, J., R. Michaely, M. Richardson, and M. R. Roberts (2007): "On the importance of measuring payout yield: Implications for empirical asset pricing," *The Journal of Finance*, 62, 877–915.

Busetti, F. and J. Marcucci (2013): "Comparing forecast accuracy: a Monte Carlo investigation," *International Journal of Forecasting*, 29, 13–27.

Campbell, J. Y. and S. B. Thompson (2008): "Predicting excess stock returns out of sample: Can anything beat the historical average?" *Review of Financial Studies*, 21, 1509–1531.

Carhart, M. M. (1997): "On persistence in mutual fund performance," *The Journal of finance*, 52, 57–82.

Clark, T. E. and K. D. West (2007): "Approximately normal tests for equal predictive accuracy in nested models," *Journal of econometrics*, 138, 291–311.

Cohen, L. and A. Frazzini (2008): "Economic links and predictable returns," *The Journal of Finance*, 63, 1977–2011.

Da, Z., J. Engelberg, and P. Gao (2011a): "In search of attention," *The Journal of Finance*, 66, 1461–1499.

——— (2011b): "In search of attention," *The Journal of Finance*, 66, 1461–1499.

Daley, B. and B. Green (2012): "Waiting for News in the Market for Lemons," *Econometrica*, 80, 1433–1504.

Engelberg, J. (2008): "Costly information processing: Evidence from earnings announcements," .

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

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

Fang, L. H., J. Peress, and L. Zheng (2014): "Does Media Coverage of Stocks Affect Mutual Funds' Trading and Performance?" *The Review of Financial Studies*, 27, 3441–3466.

George, T. J. and C.-Y. Hwang (2004): "The 52-week high and momentum investing," *The Journal of Finance*, 59, 2145–2176.

Goyal, A. and I. Welch (2008): "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies*, 21, 1455–1508.

Gurun, U. G. and A. W. Butler (2012): "Don't believe the hype: Local media slant, local advertising, and firm value," *The Journal of Finance*, 67, 561–598.

Hillert, A., H. Jacobs, and S. Müller (2014): "Media makes momentum," *The Review of Financial Studies*, 27, 3467–3501.

Hong, H., T. Lim, and J. C. Stein (2000): "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies," *The Journal of Finance*, 55, 265–295.

Hong, H. and J. C. Stein (2007): "Disagreement and the stock market," *Journal of Economic perspectives*, 21, 109–128.

Hong, H., W. Torous, and R. Valkanov (2007): "Do industries lead stock markets?" *Journal of Financial Economics*, 83, 367–396.

Huang, D., F. Jiang, J. Tu, and G. Zhou (2014): "Investor sentiment aligned: A powerful predictor of stock returns," *Review of Financial Studies*, hhu080.

Jacobs, H. (2015): "The role of attention constraints for investor behavior and economic aggregates: what have we learnt so far?" *Management Review Quarterly*, 65, 217–237.

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

Kandel, S. and R. F. Stambaugh (1996): "On the predictability of stock returns: an asset-allocation perspective," *The Journal of Finance*, 51, 385–424.

Kelly, B. and S. Pruitt (2013): "Market expectations in the cross-section of present values," *The Journal of Finance*, 68, 1721–1756.

Lettau, M. and S. Ludvigson (2001): "Consumption, aggregate wealth, and expected stock returns," *the Journal of Finance*, 56, 815–849.

Li, F. W. (2016): "Macro Disagreement and the Cross-Section of Stock Returns," *The Review of Asset Pricing Studies*, 6, 1–45.

Loughran, T. and B. McDonald (2011): "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of Finance*, 66, 35–65.

Markowitz, H. (1952): "Portfolio selection," *The journal of finance*, 7, 77–91.

Menzly, L. and O. Ozbas (2010): "Market segmentation and cross-predictability of returns," *The Journal of Finance*, 65, 1555–1580.

Miller, E. M. (1977): "Risk, uncertainty, and divergence of opinion," *The Journal of finance*, 32, 1151–1168.

Morck, R., B. Yeung, and W. Yu (2000): "The information content of stock markets: why do emerging markets have synchronous stock price movements?" *Journal of financial economics*, 58, 215–260.

Newman, M. (2010): *Networks: an introduction*, Oxford university press.

Odean, T. (1999): "Do Investors Trade Too Much?" *American Economic Review*, 89, 1279–1298.

Özgür, A., B. Cetin, and H. Bingol (2008): "Co-occurrence network of reuters news," *International Journal of Modern Physics C*, 19, 689–702.

Peng, L. and W. Xiong (2006): "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, 80, 563–602.

Peng, L., W. Xiong, and T. Bollerslev (2007): "Investor Attention and Time-varying Comovements," *European Financial Management*, 13, 394–422.

Rapach, D., J. Strauss, J. Tu, and G. Zhou (2015): "Industry interdependencies and cross-industry return predictability," *Working Paper*.

Rapach, D. E., J. K. Strauss, and G. Zhou (2010): "Out-of-sample equity premium prediction: Combination forecasts and links to the real economy," *Review of Financial Studies*, 23, 821–862.

Rubin, A., B. Segal, and D. Segal (2017): "The interpretation of unanticipated news arrival and analysts skill," *Journal of Financial and Quantitative Analysis*, 52, 1491–1518.

Scherbina, A. and B. Schlusche (2015): "Economic linkages inferred from news stories and the predictability of stock returns," *Working Paper*.

Solomon, D. H., E. Soltes, and D. Sosyura (2014): "Winners in the spotlight: Media coverage of fund holdings as a driver of flows," *Journal of Financial Economics*, 113, 53–72.

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

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

Yu, Y. (2015): "Market-wide attention, trading, and stock returns," *Journal of Financial Economics*, 116, 548–564.

Zhang, J. L., W. K. Härdle, C. Y. Chen, and E. Bommes (2016): "Distillation of news flow into analysis of stock reactions," *Journal of Business & Economic Statistics*, 34, 547–563.