News Sentiment to Market Impact and its Feedback Effect

Sheung Yin Kevin Mo, Anqi Liu and Steve Y. Yang

(Financial Engineering Program, Stevens Institute of Technology, USA)

Abstract: Although market feedback on investor sentiment effect has been conceptually identified in the existing finance literature, and investment strategies have been designed to explore this effect, there lacks systematic analysis in a quantified manner on such effect. Digitization of news articles and the advancement of computational intelligence applications have led to a growing influence of news sentiment over financial markets in recent years. News sentiment has often been used as a proxy for gauging investor sentiment and reflecting the aggregate confidence of the society toward future market. Previous studies have primarily focused on elucidating the unidirectional impact of news sentiment on market returns and not vice versa. In this study, we analyze more than 12 millions of news articles and document the presence of a significant feedback effect between news sentiment and market returns across the major indices in the U.S. financial market. More specifically, we find that news sentiment exhibits a lag-5 effect on market returns and conversely market returns elicit consistent lag-1 effects on news sentiment. This aligns well with our intuition that news sentiment drives trading activity and investment decisions. In turn, heightened investment activity further stimulates involuntary responses, which manifest in the form of more news coverage and publications. The evidence presented highlights the strong correlation between news sentiment and market returns, and demonstrates the benefits of advancing knowledge in data-driven modeling and its interaction with market movements.

Keywords: News sentiment; market returns; feedback; regression analysis

1. Introduction

The motivation of this study hinges on the growing digitization of news articles and the advancement of computational intelligence applications in analyzing news content. News has traditionally served as an important information source, affecting financial environments of all sizes from corporate levels to macroeconomics. Until recently, news information has become more frequently used by institutional investors and algorithmic trading systems (Johnson, 2010). As a consequence, there exists a wealth of *commercial-of-the-shelf* news analytics packages such as the likes of Bloomberg's News Sentiment Analysis App, Thomson Reuters's NewsScope, OptiFine, FINIF and FinSentS. Through computerized news handling techniques, news articles are converted to a quantitative measure for sentiment representation. In the field of behavioral finance, news sentiment has often been used as a proxy for gauging investor's sentiment and reflecting the aggregate confidence of the society toward financial markets (Q. Li et al.,

---

Corresponding Authors' Names and Affiliations

- Sheung Yin Kevin Mo, Adjunct Faculty, Financial Engineering program, Stevens Institute of Technology; E-mail: smo@stevens.edu.
- Anqi Liu, PhD Candidate, Financial Engineering program, Stevens Institute of Technology; E-mail: aliu@stevens.edu.
- Steve Yang, Assistant Professor, Financial Engineering program, Stevens Institute of Technology; E-mail: syang14@stevens.edu.

2014). Intuitively, news sentiment becomes a key factor in driving financial market movement, and as a result a large surge of sentiment can enhance the predictability of the market's direction.

Empirical evidence has shown the effectiveness of using news sentiment to predict market returns (Chan, 2003; Tetlock, Saar-Tsechansky, & MacKassy, 2008; Wuthrich et al., 1998). The financial market exhibits an asymmetric response to negative news sentiment in the gold futures returns (Smales, 2014b). In addition, news sentiment has been found to present potentials for developing profitable trading systems based on accurate forecasts (X. Li, Xie, Chen, Wang, & Deng, 2014; Schumaker & Chen, 2009). In other studies related to news impact on volatility, Ho et al. demonstrated that public news sentiment is related to the intraday volatility of the Dow Jones Composite Average and negative news sentiment has a larger impact on volatility than positive news (Ho, Shi, & Zhang, 2013). Another similar study suggested that changes in S&P 500 volatility level (i.e. VIX) has a negative contemporaneous relationship with news sentiments (Smales, 2014a). However, these studies primarily focus on elucidating the unidirectional impact of news sentiment on market returns and not vice versa.

Feedback mechanisms have been explored in the field of finance, mainly through the examination of its effects on price and volatility. For effect on stock prices, Hirshleifer et al. presented a theoretical framework that justifies irrational investors to earn abnormal profits based on a feedback mechanism from stock prices to cash flows (Hirshleifer, Subrahmanyam, & Titman, 2006). Khanna and Sonti further demonstrated the feedback effect of stock prices on firm value through a herding equilibrium model and investigated into the incentive for traders to conduct price manipulation (Khanna & Sonti, 2004). Crude oil prices were found to elicit feedback effects along with an inverse leverage impact with its implied volatility (Aboura & Chevallier, 2013). On the other hand, empirical evidence suggests feedback effect between squared volatility and stock price. Inkaya and Okur showed that large feedback effect rate is a useful indicator for measuring market stability by estimating volatility feedback effect rate using Malliavin (İnkaya & Yolcu Okur, 2014). There is also empirical evidence that feedback trading, a self-perpetuating pattern of investor's behavior, is present in G7 stock markets, and other international markets (Antoniou, Koutmos, & Pericli, 2005; Salm & Schuppli, 2010). Furthermore, the effect of feedback trading was found to vary across business cycle (Chau & Deesomsak, 2014) and the strongest influence was observed during periods of financial crisis with declining futures prices (Salm & Schuppli, 2010). Hou and Li developed a regression model of feedback trading to analyze CSI300 stock returns and demonstrated that lagged index returns can predict market index return and conditional volatility (Hou & Li, 2014). Feedback trading was also found to significantly influence exchange rate movements (Laopodis, 2005). Using a theoretical framework, Arnold and Brunner showed that positive feedback trading causes price overreaction and the impacts of feedback trading would be dampened if news is incorporated into price in time (Arnold & Brunner, 2014).

Our current study uses a data-driven approach to document the feedback effect between news sentiment and market returns in a quantifiable manner, as it has not been investigated in previous literature. The study advances the understanding of news sentiment in the context of behavioral finance and gain insight with its market impact, specifically in answering the following two research questions:

i. Does a contemporaneous quantitative relation exist between news sentiment and market returns?
ii. How does sentiment expressed by news articles exert feedback effect on financial market returns?

The remainder of the paper is structured as follows: Section 2 details the data collection and processing procedures for determining news sentiment. Section 3 explores the relationship between news sentiment and market returns in terms of their contemporaneous and feedback effect. Section 4 discusses the intuition behind the feedback mechanism through a narrative description and ends with the conclusions of the study in Section 5.

2. Data

2.1 News Sentiment

We collect a total of 12,842,637 news articles between July 13, 2012 and February 27, 2015 from the Northern Light SinglePoint Business News portal. Northern Light is a data vendor that sources business content across news publishers and online newswires. It is chosen as our primary source of news data because of its high relevance towards the financial market and comprehensive coverage of information. As part of the collection effort, we build a news crawler that extracts relevant market-related news entries from the portal database and pre-processes them into news sentiment (see Figure 1). The news crawler features a Java-based platform that searches relevant news related to the financial markets, and records attributes such as the title, summary, description and the sentiment. In this study, we utilize major financial market indices, S&P 500, Dow Jones and Russell 3000, as proxies for financial news sources to represent the overall market sentiment. The data contains 660 business days across the evaluation period, in which only weekday data is utilized. In addition, a total of 2,420 distinct news providers are captured in this dataset representing a diverse group of media sources. Examples of news publishers include the Wall Street Journal and the New York Times. For each news media provider, we record their respective frequency of publications on a daily basis.

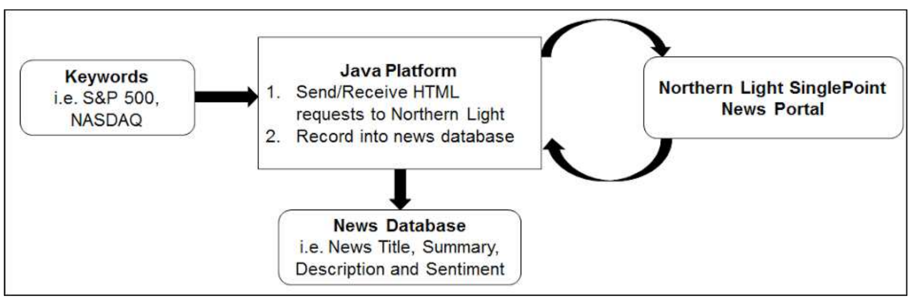

Figure 1: News Crawler from Northern Light Business News Portal

Commonly used sentiment algorithms can be categorized into two major groups: lexicon-based approach and machine learning techniques. Medhat et al. conducted an extensive survey of current sentiment analysis algorithms and found that the lexicon-based approach has been more often used in research studies from 2010 to 2013 (Medhat, Hassan, & Korashy, 2014). In addition, Moreo et al. proposed a lexicon-based news sentiment analyzer that can incorporate non-standard language and generate sentiment measures based on specific topics of interest (Moreo, Romero, Castro, & Zurita, 2012). A study conducted by Schumaker et al. investigated the effectiveness of the Arizona Financial Text system, which leverages the use of OpinionFinder for identifying the text’s tone and polarity (Schumaker, Zhang, Huang, & Chen, 2012). A lexicon-based approach, Li et al. evaluated financial news articles using the Harvard psychological dictionary and Loughran-McDonald financial sentiment dictionary for sentiment generation (X. Li et al., 2014). Other studies have proposed the approach to focus on the use of emotion words such as “soar” and “fall” to enhance the sentiment of news articles (Yu, Wu, Chang, & Chu, 2013).

For our study, we prefer the use of the lexicon-based approach and specifically utilize word dictionaries to generate news sentiment scores. Our sentiment algorithm is developed based on the use of the SentiWordNet dictionary, a lexical resource with words linked to sentimental scores (Baccianella, Esuli, & Sebastiani, 2010). Through a four-step procedure, we convert the raw text format into daily news sentiment score for the empirical study. With the complex textual structure, we initially decompose the raw text into individual words with the removal of stop words. We then apply lemmatization techniques to convert different inflicted forms of a word into a uniform entity. For instance, we would regard "rising", "risen" and "rises" as the word entity "rise". For each word in the news text, we extract the associated score from the sentiment dictionary and finally, we generate the sentiment score for each news text by averaging all individual word scores. To compute the daily news sentiment, we aggregate all news articles published in each day and compute the daily average value of news sentiment scores.

$$News_t = \frac{1}{N(t)} \sum_{i=1}^{N(t)} Sentiment(i)$$

$$Sentiment(i) = \frac{1}{n(i)} \sum_{j=1}^{n(i)} \mathrm{sgn}(j) \times Score(j)$$

$$\mathrm{sgn}(j) = \begin{cases} 1, & word\ j\ is\ in\ the\ SentiWordNet\ dictionary \\ 0, & word\ j\ is\ not\ in\ SentiWordNet\ dictionary \end{cases}$$

where $News_t$ is the daily sentiment score, $N(t)$ is the total number of news article in a day, $Sentiment(i)$ is the sentiment score for each news article, $n(i)$ is the number of words in the news article i, and $Score(j)$ is the SentiWordNet score of word j.

2.2 Financial Market Indices – SPY, DIA, QQQ, IWV

Four major financial market index ETFs are used for empirical analysis: SPDR S&P 500, SPDR Dow Jones Industrial Average, PowerShares QQQ Trust, and SPDR Russell 3000. These indices represent an accurate proxy to financial market movement. We collect daily historical prices of these indices through Bloomberg Terminal from July 13, 2012 and February 27, 2015. The respective price index is transformed into returns time series and then aligned with corresponding news sentiment values. The resultant return series represent total return indices, which assume that all cash payouts are reinvested automatically.

$$R_{I_t} = \ln \frac{I_t}{I_{t-1}}$$

where R_{I_t} is the return of the market index, I_t is the market index price at time t.

3. Relationship Between News Sentiment and Market Movement

3.1 News Sentiment Feedback Effect Hypothesis

The objective of the paper is to examine the hypothesis that there is a feedback effect between news sentiment and financial market movement. Based on our intuition that news sentiment drives trading activity and investment decisions, news sentiment should have a direct impact on market returns. Furthermore, we posit that the heightened investment activity stimulates involuntary responses, which manifest in the form of more news coverage and publications. This section introduces the conceptual roadmap from a theoretical perspective of how news sentiment behaves and interacts with market movement (see Figure 3). It is important to note that the feedback mechanism is only effective in specific time dimensions, i.e. different time lags in each direction. In this paper, we conduct empirical tests to validate the hypothesis of the news sentiment feedback effect in the form of regression models, vector autoregression model and the Granger Causality test. We seek to explore the fundamental relationships between sentiments expressed by news against major market index returns. The top chart of Figure 4 shows the three time series of sentiment score for S&P 500, Dow Jones and NASDAQ respectively, while the bottom chart shows the return series of the S&P 500 index for illustration purpose.

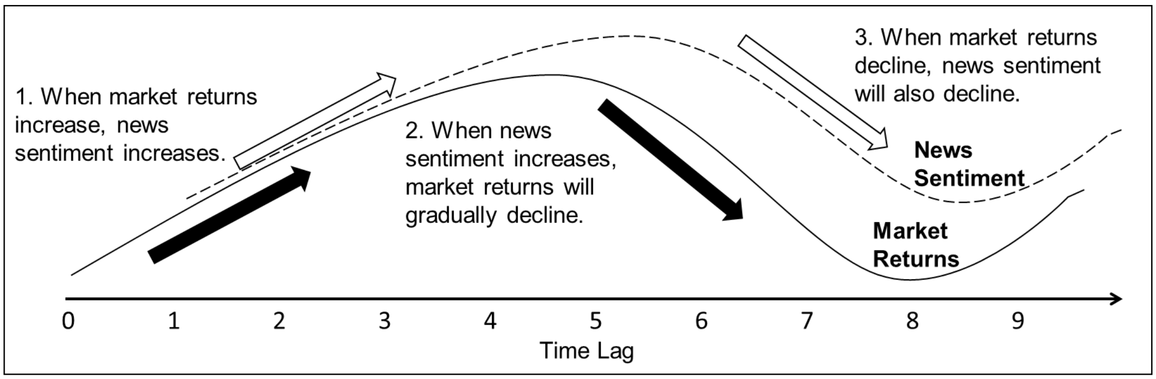

Figure 2: News Sentiment and its Feedback Effect

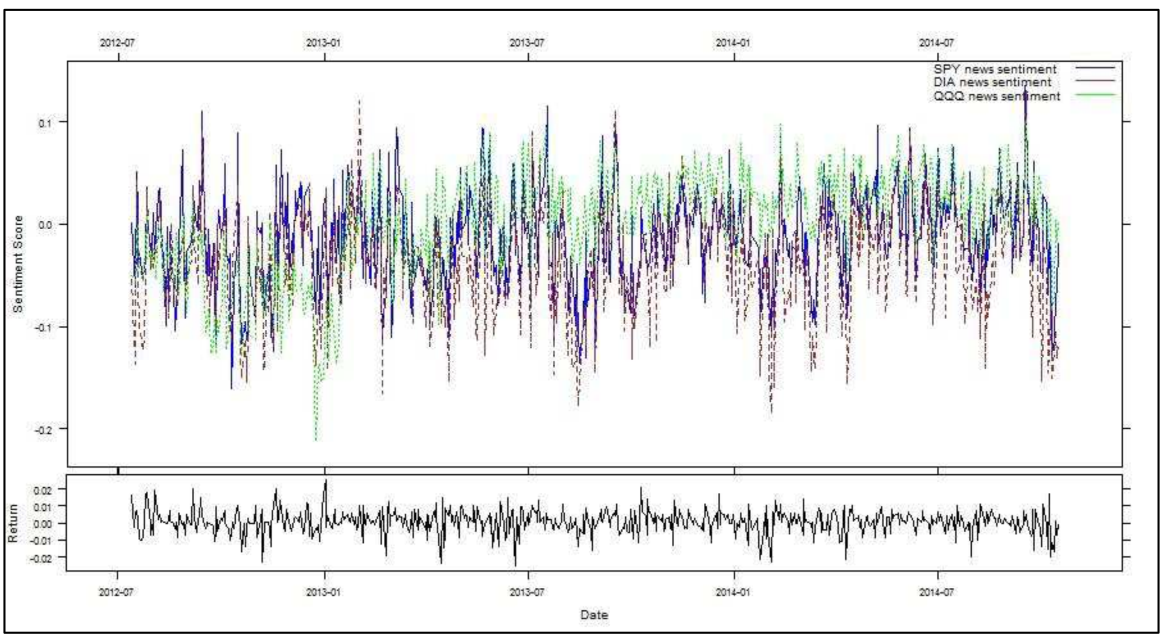

Figure 3: News Sentiment and Market Movement

3.2 The Contemporaneous Relationship between News Sentiment and Market Movement

News sentiment is a key factor in driving market movement and we hypothesize that market movement also has an effect on news sentiment. For preliminary investigation, we are interested in determining the contemporaneous relationship between news sentiment and the market movement. We first establish linear model with news sentiment as explanatory variables to S&P 500 return.

$$R_t = \beta_0 + \beta_1 News_t + \epsilon_t$$

where R_t is S&P 500 index return on day t, $News_t$ is news sentiment on day t, and ϵ_t is the residual error.

Using the ordinary least squares method, we find that the linear model is significant in its respective explanatory variable suggesting the strong relationship between news sentiment and market movement (see Table 1). From the model result, news sentiment is found to have a significant relationship with the S&P 500 market index with a positive coefficient of 0.0578. The result suggests that news sentiment has a positive influence on the same day towards market returns.

$$R_t = 0.0026 + 0.0578 News_t$$

Table 1: Contemporaneous Relationship between News Sentiment and SPY Return

Lag-0 Regression	News Sentiment -> S&P 500 Return
Number of Observation	660
Coefficient	0.0578
T-statistics	11.843
p-value	**<2.0e-16*** **

3.3 Linear Regression across Different Time Lags

With the evidence of contemporaneous relationship, we investigate the feedback effect across different time lags between news sentiment and market movement. The objective is to understand the bilateral relations between the two time series at varying time lags. For example, what is the effect of market movement on the news sentiment after a time lag of several days? We are specifically interested in understanding the lagged effect of one variable on another. By doing so, we construct a similar linear model structure with the lagged value of the explanatory variable. Since our data is aggregated on a daily basis, one lag is equivalent to the duration of one day.

$$R_t = \beta_0 + \beta_1 News_{t-i} + \epsilon_t$$
$$News_t = \tilde{\beta}_0 + \tilde{\beta}_1 R_{t-i} + \tilde{\epsilon}_t$$

where R_t is market index returns on day t, $News_t$ is news sentiment t, and ϵ_t and $\tilde{\epsilon}_t$ are unexplained model component.

The cross-sectional regression results are displayed in terms of their coefficients and p-values (see Table 2 to Table 5). The positive sign of the coefficient indicates that one variable has a positive impact on the other and vice versa. For p-values, we consider regression result less than 0.05 to be significant. We find that the news sentiment at lag-5 has significant effects on all market returns for S&P 500, Dow Jones, NASDAQ and Russell 3000, suggesting that the news sentiment from five days ago has a fundamental relationship with the current market movement. For instance, the regression model for the S&P 500 index reflects a significant delayed impact with a negative coefficient of -0.0147 associated with a low p-value of 0.006**. On the other hand, we find that market movement has a more immediate and pronounced effect on news sentiment. The results show that lag-1 to lag-3 of all market returns, with the exception of NASDAQ, are positively related to the current news sentiment with significant p-values with the largest coefficient at lag-1 followed by lag-2 and lag-3. NASDAQ shows resemblance of similar result for lag-1 and lag-2 but its p-value at lag-3 is 0.066 which is not within the significance level. The discrepancy could be due to the noise for the wide coverage of the underlying market with over 3,000 companies. After all, the results illustrate the consistent and similar pattern across market index returns and together assemble evidence of a feedback relationship between news sentiment and market return, where both variables are significantly related to one another with intuitive time effect.

Table 2: Linear Regression Models between News Sentiment and SPY Return (Lag-1 to Lag-5)

	News Sentiment -> SPY Return		SPY Return -> News Sentiment	
	Coefficient	p-value	Coefficient	p-value
Lag-1	0.0028	0.599	**2.3296**	**<2.0e-16*****
Lag-2	-0.0073	0.173	**0.7836**	**0.006****
Lag-3	-0.0071	0.188	**0.7163**	**0.011***
Lag-4	**-0.0129**	**0.017***	-0.1621	0.568
Lag-5	**-0.0147**	**0.006****	0.0008	0.998

Table 3: Linear Regression Models between News Sentiment and DIA Return (Lag-1 to Lag-5)

	News Sentiment -> DIA Return		DIA Return -> News Sentiment	
	Coefficient	p-value	Coefficient	p-value
Lag-1	0.0050	0.326	**2.5845**	**<2.0e-16*****
Lag-2	-0.0033	0.517	**0.8929**	**0.003****
Lag-3	-0.0023	0.648	**0.8956**	**0.003****
Lag-4	-0.0091	0.074	-0.0654	0.828
Lag-5	**-0.0121**	**0.02***	-0.0487	0.872

Table 4: Linear Regression Models between News Sentiment and QQQ Return (Lag-1 to Lag-5)

	News Sentiment -> QQQ Return		QQQ Return -> News Sentiment	
	Coefficient	p-value	Coefficient	p-value
Lag-1	0.0049	0.433	**1.9044**	**7.18e-16*****
Lag-2	-0.0093	0.138	**0.6457**	**0.007****
Lag-3	-0.0101	0.110	0.4453	0.066
Lag-4	**-0.0132**	**0.037***	-0.2169	0.371
Lag-5	**-0.0169**	**0.007****	0.1153	0.634

Table 5: Linear Regression Models between News Sentiment and IWV Return (Lag-1 to Lag-5)

	News Sentiment -> IWV Return		IWV Return -> News Sentiment	
	Coefficient	p-value	Coefficient	p-value
Lag-1	0.0022	0.694	**2.3188**	**<2.0e-16*****
Lag-2	-0.0079	0.149	**0.7865**	**0.004****
Lag-3	-0.0079	0.152	**0.6651**	**0.016***
Lag-4	-0.0137	0.013	-0.2022	0.466
Lag-5	**-0.0152**	**0.006****	0.0229	0.934

3.4 Vector Autoregression (VAR) Model

This section aims to further explore the time series interdependencies between market index returns and news sentiment through Vector Autoregression (VAR). First, the Dickey-Fuller test is applied to validate that the time series of the market index returns and news sentiment are stationary. For robustness check, three versions of the Dickey-Fuller test are implemented, all resulting in the rejection of the null hypothesis due to significant low p-values. These desirable results confirm that the underlying time series has no unit root (See Table 6), and therefore suggest that the time series in our experiments is stationary.

$\mathbf{Dickey-Fuller\ Test}$:
Test unit root: $\nabla y_t = \delta y_{t-1} + u_t$
Test unit root with drift: $\nabla y_t = a_0 + \delta y_{t-1} + u_t$
Test unit root with drift and time trend: $\nabla y_t = a_0 + a_1 t + \delta y_{t-1} + u_t$

Table 6: Dickey-Fuller Test

Methods		*SPY Return*	*DIA Return*	*QQQ Return*	*IWV Return*	*News Sentiment*
None	t-statistics	-24.33	-24.25	-22.95	-24.12	-14.44
	p-value	<2e-16***	<2e-16***	<2e-16***	<0.01**	<2e-16***
Drift	t-statistics	-24.515	-24.369	-23.087	-24.080	-15.885
	p-value	<2e-16***	<2e-16***	<2e-16***	<2e-16***	< 2e-16***
Trend	t-statistics	-24.548	-24.409	-23.074	-24.12	-16.468
	p-value	<2e-16***	<2e-16***	<2e-16***	<2e-16***	< 2e-16***

The previous section shows statistical evidence that lagged news sentiment is a significant explanatory factor to market returns. With the strong indication that the lagged effect can last up to at least five days, we establish the vector autoregressive model to investigate the linear interdependencies among news sentiment and market returns. Using the AIC model selection criterion, we select a lag-5 model for each market index. The selection process starts with lag-10, and we find lag-5 model gives the most AIC reduction. Similar to the linear regression model, we set the confidence level to 95% in the VAR(5) model. Table 7 displays the results with the significant relation in their respective coefficients, p-values and lagged effects. The results confirm that lag-5 news sentiment has a consistent influence on all market returns. On the other hand, lag-1 and lag-4 market index returns, and lag-1 news sentiment are stable factors in this analysis. These findings align with the previous results from the linear regression model and they further confirm the significant feedback effect between market and news sentiment.

$$\mathbf{VAR(5)Model}:$$
$$News_t = \alpha_1 + A_{1,1}News_{t-1} + A_{1,2}R_{t-1} + A_{2,1}News_{t-2} + A_{2,2}R_{t-2} + \cdots + A_{5,2}R_{t-5} + \epsilon_{1,t}$$
$$R_t = \alpha_2 + B_{1,1}News_{t-1} + B_{1,2}R_{t-1} + B_{2,1}News_{t-2} + B_{2,2}R_{t-2} + \cdots + B_{5,2}R_{t-5} + \epsilon_{2,t}$$

where R_t is the market index return at time t, $A_{i,1}$ is coefficient of lag-i news to sentiment, $A_{i,2}$ is coefficient of lag-i return to sentiment, $B_{i,1}$ is coefficient of lag-i news to return, $B_{i,2}$ is coefficient of lag-i return to return, $News_t$ is news sentiment at time t, and $\epsilon_{1,t}$ and $\epsilon_{2,t}$ are white noises.

As an extension of the analysis, we find that the time series of market returns contains no autocorrelation which does not override the impact from news sentiment. This is an indication that news, an exogenous information source, is a leading factor of market returns and the negative coefficient shows the correction of investors' overreaction to the information after one week (see Figure 5). Furthermore, we examine the predictive power of information with higher lags in justifying the model selection process. We evaluate the statistical relationships between higher-lag news sentiment and market returns, aiming specifically to test if the lag-5 news sentiment contains most of the information associated with the subsequent market return changes. Our findings show that news sentiment with higher lags beyond lag-5 does not add improvement in significance and therefore validate the impact of the lag-5 effect.

Table 7: Vector Autoregressive Model (up to lag-5)

Relation	*Factor*	*Coefficient*	*t-value*	*p-value*
News, SPY Return -> SPY Return	**Lag-5 News**	-0.01	-2.19	0.029**
News, SPY Return -> News	Lag-1 SPY	1.40	4.86	1.46e-06***
	Lag-1 News	0.34	7.57	1.32e-13***
	Lag-4 SPY	-0.69	-2.35	0.019**
News, QQQ Return -> QQQ Return	Lag-2 QQQ	-0.09	-1.98	0.048**
	Lag-5 News	-0.02	-2.30	0.022**
News, QQQ Return -> News	Lag-1 QQQ	1.09	4.48	8.76e-06***
	Lag-1 News	0.36	8.08	3.12e-15***
	Lag-4 QQQ	-0.56	-2.29	0.022**
News, DIA Return -> DIA Return	**Lag-5 News**	-0.01	-2.10	0.036**
News, DIA Return -> News	Lag-1 DIA	1.57	5.12	4.07e-07***
	Lag-1 News	0.33	7.34	6.60e-13***
	Lag-4 DIA	-0.71	-2.27	0.023**
News, IWV Return -> IWV Return	**Lag-5 News**	-0.015	2.23	0.026**
News, IWV Return -> News	Lag-1 IWV	1.43	5.10	4.45e-07***
	Lag-1 News	0.34	7.53	1.75e-13***
	Lag-4 IWV	-0.70	-2.46	0.014**

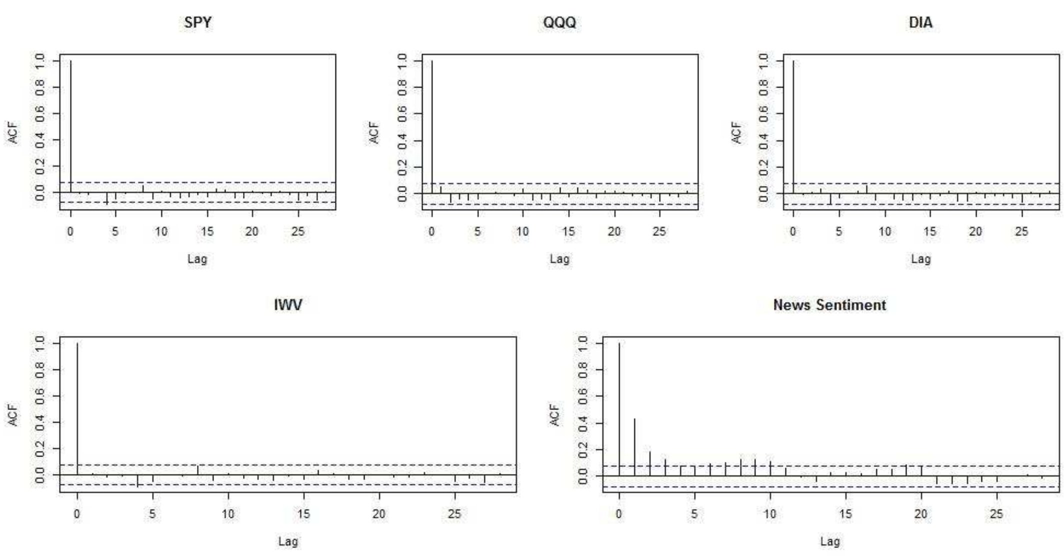

Figure 4: Autocorrelation Function of Market Returns and News Sentiment Time Series

3.5 Granger Causality Test

Granger Causality test is applied to examine the forecasting ability between news sentiment and market index returns. In the experiment, we consider significant relations that have low p-value of less than 0.10 for validation. The results from the Granger Causality test show consistency with previous findings from the linear regression and vector autoregression models. First, there exists strong statistical relation between lag-1 returns and news sentiment. The other observation suggests that the lag-5 news sentiment is correlated to major index returns. As an extended validation, we also investigate the impact of higher lags between news sentiment and market returns. When combined with the effect from lag-5 sentiment, we find that lag-6 and higher-lag sentiment do not show significant impact to the current return. This is also evident from both the VAR model and linear regression models.

Table 8: Granger Causality Test

Relation	*Lag*	*F-value*	*p-value*
SPY Returns -> News	Up to 1	19.71	1.06e-05***
News -> SPY Returns	Up to 5	1.94	0.086*
QQQ Returns -> News	Up to 1	17.38	3.48e-05***
News -> QQQ Returns	Up to 5	1.92	0.090*
DIA Returns -> News	Up to 1	22.35	2.79e-06***
News -> DIA Returns	Up to 5	1.72	0.127
IWV Returns -> News	Up to 1	21.95	3.40e-06***
News -> IWV Returns	Up to 5	1.88	0.096*

4. Feedback Discussion

The bilateral findings demonstrate that the news sentiment and market returns have a fundamental feedback relationship. This study confirms this statistical relationship through the ordinary linear regression, vector autoregression and Granger Causality test respectively with high degree of confidence level. Varying across different time lags, the study confirms that news sentiment has a significant negative relation with the market returns at lag-5 while market returns also have a more short-term impact towards news sentiment at lag-1. This result is encouraging as the presence of feedback effect in news sentiment is a novel finding compared to existing literature which has often emphasized the single direction of the news sentiment impact (Q. Li et al., 2014; X. Li et al., 2014; Smales, 2014b). It also sheds light on the formation of news sentiment and how it interacts with the market movement.

Another observation involves the chosen period of study from 2012 to 2015, which signifies the recovery period from the 2008 financial crisis. Despite the strong bull market, our methodology has shown to produce consistent results across different market conditions within the period. With the market being relatively stable in 2013 and 2014, the news sentiment feedback effect yields similar findings in each individual year within the period of study. The key findings from the linear regression model, the VAR model and the Granger Causality test further validate the presence of the sentiment feedback effect, which avoids the excess dependency of relying on an individual model. Comparing across the four major indices, we find that the results show a high level of consistency and similarity in terms of their significance levels and coefficient signs. From the results of the linear regression and VAR models, the strong lag-5 impact of news sentiment on market returns is consistent across the four market indices with significant levels. The sign of the coefficient indicates that the news sentiment from 5 days ago induces an opposite reaction of market returns suggesting a correction of the potential overreaction to the news information one week ago. The result aligns with the finding by Schumaker et al. that positive sentiment predicts price decreases and negative sentiment predicts price increases, exhibiting investors' contrarian behaviors in their decisions (Schumaker et al., 2012). On the other hand, the lag-1 positive correlation from market returns to news sentiment is the strongest by significance level. This suggests that market returns has a significant positive impact on news sentiment. In addition, the Granger Causality test illustrates the significant forecasting power of lag-1 market returns to news sentiment and up to lag-5 news sentiment to returns.

5. Conclusion

This paper presents the evidence that there exists a feedback mechanism between news sentiment and market returns among the major U.S. financial market indices, namely S&P 500, NASDAQ, Dow Jones Industrial Average, and Russell 3000. Our analysis shows that news sentiment exhibits a lag-5 effect on market returns and conversely, market returns elicit consistent lag-1 effects on news sentiment. Based on these results, we suggest that the aggregate sentiment from news articles has a more delayed impact on market returns than that of market returns on news sentiment within this feedback cycle. This aligns well with our intuition that news sentiment, a proxy for measuring aggregate confidence of the society toward financial market futures, drives trading activity and investment decisions. In turn, heightened investment activity further stimulates an involuntary response, which manifests in the form of more news coverage and publications. The evidence presented highlights the strong correlation between news sentiment and market returns, and demonstrates the benefits of advancing knowledge in data-driven modeling. For future work, we seek to expand the current investigation of time lag effects onto intraday data. We posit that pronounced differences may exist in the feedback mechanisms at different time scales, which may have potential applications in identifying profitable trading opportunities and portfolio diversification. Second, the study can benefit from the exploration of developing a trading indicator based on the feedback mechanism of news sentiment versus market returns. The indicator can be evaluated as a potential factor of an investment strategy.

References:

Aboura, S., & Chevallier, J. (2013). Leverage vs. feedback: Which Effect drives the oil market? *Finance Research Letters*, *10*(3), 131–141.

Antoniou, A., Koutmos, G., & Pericli, A. (2005). Index futures and positive feedback trading: evidence from major stock exchanges. *Journal of Empirical Finance*, *12*(2), 219–238. http://doi.org/10.1016/j.jempfin.2003.11.003

Arnold, L. G., & Brunner, S. (2014). The economics of rational speculation in the presence of positive feedback trading. *The Quarterly Review of Economics and Finance*.

Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, … D. Tapias (Eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)* (pp. 2200–2204). European Language Resources Association (ELRA).

Chan, W. S. (2003). Stock price reaction to news and no-news: drift and reversal after headlines. *Journal of Financial Economics*, *70*(2), 223–260. http://doi.org/10.1016/S0304-405X(03)00146-6

Chau, F., & Deesomsak, R. (2014). Business cycle variation in positive feedback trading: Evidence from the G-7 economies. *Journal of International Financial Markets, Institutions and Money*.

Hirshleifer, D., Subrahmanyam, A., & Titman, S. (2006). Feedback and the success of irrational investors. *Journal of Financial Economics*, *81*(2), 311–338.

Ho, K.-Y., Shi, Y., & Zhang, Z. (2013). How does news sentiment impact asset volatility? Evidence from long memory and regime-switching approaches. *The North American Journal of Economics and Finance*, *26*, 436–456.

Hou, Y., & Li, S. (2014). The impact of the CSI 300 stock index futures: Positive feedback trading and autocorrelation of stock returns. *International Review of Economics & Finance*, *33*, 319–337.

İnkaya, A., & Yolcu Okur, Y. (2014). Analysis of volatility feedback and leverage effects on the ISE30 index using high frequency data. *Journal of Computational and Applied Mathematics*, *259*, 377–384.

Johnson, B. (2010). *Algorithmic Trading & DMA: An Introduction to Direct Access Trading Strategies* (p. 574). 4Myeloma Press.

Khanna, N., & Sonti, R. (2004). Value creating stock manipulation: feedback effect of stock prices on firm value. *Journal of Financial Markets*, *7*(3), 237–270.

Laopodis, N. T. (2005). Feedback trading and autocorrelation interactions in the foreign exchange market: Further evidence. *Economic Modelling*, *22*(5), 811–827.

Li, Q., Wang, T., Li, P., Liu, L., Gong, Q., & Chen, Y. (2014). The effect of news and public mood on stock movements. *Information Sciences*, *278*, 826–840.

Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, *69*, 14–23.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113.

Moreo, A., Romero, M., Castro, J. L., & Zurita, J. M. (2012). Lexicon-based Comments-oriented News Sentiment Analyzer system. *Expert Systems with Applications*, *39*(10), 9166–9180.

Salm, C. a., & Schuppli, M. (2010). Positive feedback trading in stock index futures: International evidence. *International Review of Financial Analysis*, *19*(5), 313–322.

Schumaker, R. P., & Chen, H. (2009). A quantitative stock prediction system based on financial news. *Information Processing & Management*, *45*(5), 571–583.

Schumaker, R. P., Zhang, Y., Huang, C.-N., & Chen, H. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems*, *53*(3), 458–464.

Smales, L. (2014a). News sentiment and the investor fear gauge. *Finance Research Letters*, *11*(2), 122–130. http://doi.org/10.1016/j.frl.2013.07.003

Smales, L. (2014b). News sentiment in the gold futures market. *Journal of Banking & Finance*, *49*, 275–286.

Tetlock, P. C., Saar-Tsechansky, M., & MacKassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, *63*(3), 1437–1467.

Wuthrich, B., Cho, V., Leung, S., Permunetilleke, D., Sankaran, K., & Zhang, J. (1998). Daily stock market forecast from textual web data. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.98CH36218)* (Vol. 3, pp. 2720–2725). IEEE.

Yu, L.-C., Wu, J.-L., Chang, P.-C., & Chu, H.-S. (2013). Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. *Knowledge-Based Systems*, *41*, 89–97. http://doi.org/10.1016/j.knosys.2013.01.001