TWITTER AND FINANCIAL MARKETS

Muktamala Chakrabarti, Asim Kumar Pal, Ashok Banerjee
Indian Institute of Management Calcutta
D.H. Road, Joka, Kolkata, India

Abstract

Several researchers (Mitra and Mitra 2011, Tetlock 2007) have already shown that media influences investor sentiment, hence asset prices, asset price volatility and liquidity. Since 1960's researchers have been investigating the relationship between returns and press releases, M&A announcements, annual reports etc. Studies in behavioral finance (Lucey and Dowling 2005, Baker and Nofsinger 2002) have already found that emotions affect investor decisions. Social media is an important source of information flow which has tremendously grown in importance over the past few years. Although the relationship between news and financial markets has already been researched a lot, the research about the relationship between social media and financial markets is largely unexplored. In this paper we make an attempt to capture this emotion or sentiment about the company through an analysis of public tweets obtained from Twitter. We investigate the impact that the sentiments extracted from twitter postings about a company have on its performance in the markets, namely stock returns, volatility and liquidity.

Keywords: twitter; stock returns; liquidity; volatility.

I. INTRODUCTION

Stock market prediction is an extremely difficult problem and has been a long term dream for any investor or money manager. Modeling the dynamics of financial markets is an extremely complex task. Behavioral economists have examined the Efficient Market Hypothesis (EMH) thoroughly. There are three forms of EMH (Fama 1965)- weak, semi strong and strong. If the weak form of EMH holds true stock prices cannot be predicted using past prices. The semi strong form says prices include all publicly available information and if strong form of EMH holds true, prices instantly reflect all hidden or insider information also. But numerous studies have also shown that EMH does not hold true always and stock market prices can be predicted to some extent. Literature in Artificial Intelligence have tried to predict stock prices using previous stock prices with various AI techniques like neural networks, fuzzy systems etc. (Refenes 1994). There are studies (Mitra and Mitra 2011, Leinweber and Sisk 2011, Tetlock 2007, Schumaker and Chen 2009) showing the effect of news stories on stock prices. There are also studies showing the effect of other sources of information flow like social media on stock prices (Das et al. 2005, Tumarkin and Whitelaw 2001). Social media especially Facebook and Twitter have tremendously grown in importance over the years. Social media is considered to be the fastest ways of news transmission which spreads messages across the globe instantly.

Companies have started using social media analytics to understand customer sentiment in order to decide on marketing strategies. Asur and Huberman (2010) have shown how social media content can be used to predict real world outcomes like box office revenues for movies. Exploring a relationship between social media and financial markets is gradually gaining interest among researchers. Antweiler and Frank (2004) and Das and Chen (2007) have made significant contribution into this area of research. The former work has used messages from Yahoo Finance and Raging Bull to show that messages weakly predict volatility but cannot predict returns. They show that agreement among posted messages is associated with decreased trading volume. The latter work found that returns drive message board sentiment and not the other way, and the results are stronger when messages for many stocks are integrated into an index level sentiment measure. Gilbert and Karahalios (2010) have demonstrated that estimating emotions from weblogs provides novel information i.e. information not apparent in market data about future stock market prices.

Twitter is one of the most popular social networking sites available. Since its inception in 2006 it has immensely grown in popularity. "*As of December 2014, Twitter has more than 500 million users, out of which more than 284 million are active users"* (http://en.wikipedia.org/wiki/Twitter). The huge popularity of Twitter has attracted researchers from various areas. Twitter tweets have been used to understand customer insight (Chamlertwat et al. 2012). Various researchers have explored the relationship between Twitter sentiment and financial markets. Bollen et al. (2011) investigated whether public mood can be used for predicting economic indicators. They measured collective mood states from Twitter feeds and calculated its correlation with Dow Jones Industrial Average (DJIA) over time. Their results indicate that inclusion of specific public mood dimensions significantly improves the prediction accuracy of the up and down movements of DJIA. Collective hope and fear on each day has been measured from Twitter feeds in Zhang et al. (2011) and the authors found significant negative correlation between emotional tweet percentage and Dow Jones, NASDAQ and S&P 500, and significant positive correlation with VIX. VIX is the volatility index started by Chicago Board of Options Exchange. Si et al. (2013) used topic based sentiments from twitter to predict S&P 100 index.

Although the studies mentioned above explores the relationship between Twitter and financial markets none of them really look into company specific tweets and its effect on company specific returns. The studies are mostly concerned on the effect on the index. Also none of the studies mentioned above look into emerging markets. The studies include data mostly from developed markets. The characteristics of developed markets and emerging markets are different as information flow is much faster in developed markets. Here in this paper we use company specific tweets on Indian companies and explore its relationship with company specific return, liquidity and volatility. Also this study looks into the effect of tweets coming during trading hours and non-trading hours separately.

II. DATA

The data is selected for 42 companies listed in the Nifty Index. The list of the companies is presented in Table I. The traditional Information Retrieval technique using keyword search is used here for relevance filtering. The keywords used for searching each company specific tweets are given in Table I. The keywords are mostly the commonly used short forms of the company names. The assumption here is that people often use these short forms instead of writing the full names of the companies in tweets where there is a restriction of 140 characters on the length of the tweets. For some companies like Infosys more than one keyword is used because both the keywords are commonly used (see Table I). Tweets for these companies where the geographical location of the user posting that tweet is within 900 miles from 22^{o} North latitude and 82^{o} East longitude are downloaded. This geographical location is set to extract tweets mostly from Indian locations. The language of the tweets considered is restricted to English only. The searchTwitter() command from R has been used and the tweets downloaded are subjected to the restrictions imposed on this command. The time period of the data is from 11th November 2012 to 24th May 2013 (about six months). This method can be only applied to those companies whose names don"t have any meaning as an English word. For example the keyword "axis" cannot be used for the company Axis Bank because all tweets containing the word "axis" are going to be retrieved but many of them would not be relevant to the company.

III. EXPERIMENTAL RESULTS

A. Twitter sentiment and stock returns

The first experiment tries to explore the relationship between Twitter sentiment and stock returns. The number of positive and negative words is calculated for each tweet. The list of positive words and negative words are used which are same as the ones used in Hu and Liu (2004). There are some misspelled words in this list intentionally included by Hu and Liu (2004) as these misspelled words appear frequently in social media content. Some positive words and some negative words were added to that list to include some terms from the finance domain. The positive words added were "buy", "up", "upgrade", "upgraded" and "upgrading". The negative words added were "cut", "cuts", "down", "downgrades", "downgraded", "downgrading" and "sell". Each tweet is assigned a score which is equal to the number of positive words minus the number of negative words appearing in that tweet.

Two windows are created - trading window and non trading window. This segregation between the trading hour and non trading hour is made to understand the effects of trading hour and non trading hour tweets separately. The non-trading window for „today" starts from 4.00 pm IST yesterday and ends at 9.30 am IST today. If the previous day is not a trading day then the window starts from 4.00 pm IST of last trading day. The trading window for today starts at 9.30 am IST today to 4.00 pm IST today. Two sentiment scores are obtained for each day summing up the scores from tweets over the non trading window and the trading window respectively.

The effect of the tweets coming up over the non trading hours is seen on that day"s opening price and the effect of tweets coming up over the trading hours is observed on that day"s closing price for that particular stock. Table I shows the average, minimum and maximum of the number of trading and non trading hour tweets retrieved for each company during the period. Our results also corroborate the popular belief that people write more in social media during non trading hours. The average number of tweets is more during the non trading hours for most of the companies (see Table I). But it is also to be noted here that the non trading hour window is much larger than the trading hour window.

Two correlation measures are calculated for each company. The correlation measures along with the t statistics in brackets are given in Table II in Appendix.

1) CorrelationI- The correlation between the sentiment scores of trading hour tweets and excess stock returns during the trading hours. Excess stock return for trading hour for the day t is calculated as

$$\left(\frac{\text{company closing price}_t}{\text{company opening price}_t}\right) \square \ln\left(\frac{\text{Nifty closing price}_t}{\text{Nifty opening price}_t}\right) \quad (1)$$

2) Correlation II- The correlation between sentiment scores of non trading hour tweets and excess stock returns for non trading hours. Excess stock return for non trading hour for the day t is calculated as

$$\ln\left(\frac{\text{company opening price}_t}{\text{company closing price}_{t-1}}\right) \square \ln\left(\frac{\text{Nifty opening price}_t}{\text{Nifty closing price}_{t-1}}\right) \quad (2)$$

The correlation measures indicate that the tweets which come up during the trading hours are more relevant tweets as the correlation coefficient is significant for more number of companies. The paired t test was performed between Correlation I and Correlation II for all the companies and the t statistics was found to be 2.294, significant at 5 % level. It seems tweets coming up during trading hours are much more useful than which come up during non trading hours. Although people may write more during non trading hours important tweets come up more during trading hours. The figure below shows the sentiment score for trading hour from Twitter and excess return for trading hour scaled by 1000 for State Bank of India(SBI) during the period 11th November 2012 to 24th May 2013. It shows that the excess return series and sentiment score series are moving together to some extent, which implies that sentiment scores have some predictive content.

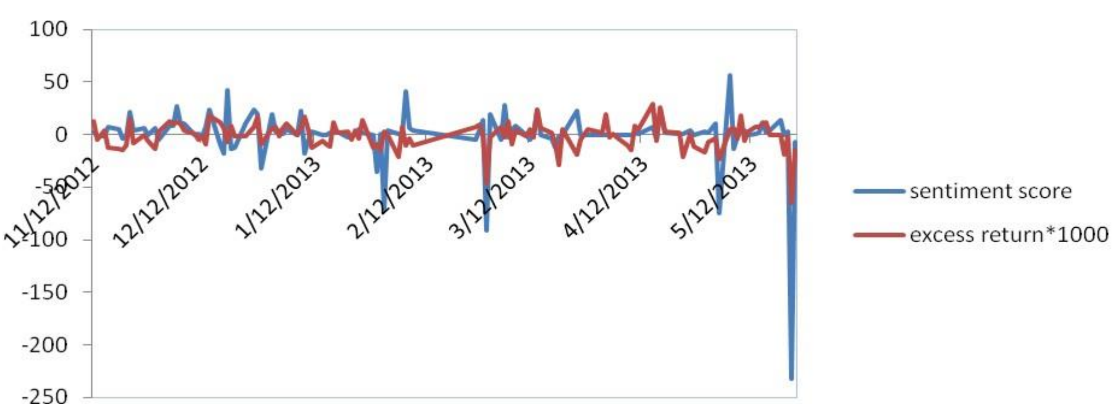

Figure 1. Sentiment Score from Twitter and excess return for SBI for the period 11th November 2012 to 24th May 2013

B. Twitter and Liquidity

A market is considered liquid if traders can trade without effecting prices significantly (Mitra et al. 2011). Liquidity is much easier defined than measured quantitatively. Trading volume is a crude measure of liquidity. A better measure of illiquidity, the Amihud illiquidity measure (Amihud 2002) is mentioned in Equation 3.The Amihud illiquidity measure is the daily ratio of absolute stock return to its dollar volume averaged over a time period.

$$ILLIQ_{iy} = 1/D_{iy} \sum_{t=1}^{D_{iy}} |R_{iyd}| / VOLD_{iyd} \qquad (3)$$

D_{iy} is the number of days for which data is available for stock i in year y, R_{iyd} is the return on stock i on day d of year y, $VOLD_{iyd}$ is the volume of stock i on day d of year y in dollars.

This Amihud illiquidity measure for annual average is transformed into a daily score by using high frequency data. Here the daily ratio of absolute stock return to its dollar volume averaged over a period is replaced by per minute interval absolute stock return to its dollar volume in that time period averaged over a full day. So we have a Amihud illiquidity score for each company for each trading day. Here in Indian context all the calculations are done in terms of Rupees.

Two correlation measures are calculated for each company. The correlation measures along with the t statistics are given in Table II-

1) CorrelationIII- The correlation between the number of tweets coming up during the trading hours and the trading volume on that day.
2) CorrelationIV- The correlation between the ln (Amihud illiquidity score) and ln (number of trading hour tweets).

The results show a strong positive relation between the number of tweets and trading volume. Almost for all the companies the coefficients (Correlation III) are statistically significant and positive. The Correlation IV is mostly negative for most of the companies. The two correlation measures show that number of tweets is a good indicator of a stock‟s liquidity. When the stock is in the news or people are talking more about that company the liquidity is also high on those days. The results clearly show that if a company is much talked about in the social media more number of people are actually interested in it which increases its trading volume and liquidity.

C. Twitter and volatility

In the next experiment the relationship between sentiment of Twitter and volatility is explored. The realized volatility is taken as a measure of volatility for this experiment. The realized volatility (squared root of the sum of the squared observed inter-period returns) is taken as a measure of volatility for this experiment. The correlation measure (Correlation V in Table II) between realized volatility and number of tweets obtained during trading hours is calculated for each company. The results show statistically significant relationship for many companies which means when the company is much talked about in the social media the volatility is also high on those days

TABLE I. DESCRIPTIVE STATISTICS

Company Name	Keyword	trading hour tweets			non trading hour tweets		
		average	max	min	average	max	min
Ambuja Cements Limited	ambuja cements	3.53	58	0	3.43	25	0
Asian Paints	asian paints	3.87	60	0	5.25	60	0
Ltd. Bajaj Auto	bajaj auto	5.89	89	0	5.42	42	0
Ltd Bank of	bank of baroda	4.07	83	0	7.25	217	0
Baroda	bhel	12.97	85	0	20.98	169	0
Bharat Heavy Electricals Ltd.	bpcl	5.36	63	0	4.87	47	0
Bharat Petroleum Corporation	airtel	278.38	582	1	553.10	1760	8
Ltd. Bharti Airtel Ltd.	cairn	7.43	119	0	8.41	156	0
Cairn India	cipla	5.75	103	0	8.06	73	0
Ltd. Cipla Ltd.	coal india reddys laboratories,	29.42	385	0	54.31	460	0
Coal India Ltd.	reddy's lab						
Dr. Reddy's Laboratories	gail	0.89	12	0	1.28	68	0
Ltd. Gail (India) Ltd.	grasim	8.34	50	0	33.80	165	0
Grasim Industries Ltd.	hcl	1.10	12	0	1.63	14	0
HCL Technologies	hdfc	28.25	575	1	41.57	286	0
Ltd. HDFC Bank Ltd.	hero motor	30.34	303	1	36.36	262	0
Hero Motocorp Ltd.	hindalco	2.07	18	0	4.65	32	0
Hindalco Industries	hul icici	3.85	28	0	3.63	48	0
Ltd. Hindustan	infy,infos	14.43	449	0	15.94	357	0
Unilever Ltd. ICICI	ys idfc	33.54	197	1	54.11	917	1
bank	itc	76.32	1095	1	129.79	1098	0
Infosys	jaiprakash	5.46	36	0	7.51	55	0
IDFCLt	associates jindal	37.34	163	0	69.73	311	0
d. ITC	steel	1.45	24	0	1.15	10	0
Ltd.	kotak mahindra	3.07	22	0	3.84	20	0
Jaiprakash Associates	lupin	3.06	28	0	4.07	41	0
Ltd. Jindal Steel Power	maruti	5.26	60	0	10.16	111	0
Ltd. Kotak Mahindra	ntpc	58.66	298	0	84.05	950	0
Bank Lupin Ltd.	ongc	11.63	299	0	17.92	208	0
Maruti Suzuki India	power grid	18.09	131	0	32.44	168	0
Ltd. NTPC Ltd.	corporation punjab	0.61	8	0	1.02	16	0
Oil and Natural Gas Corporation	national bank	2.71	96	0	4.03	211	0
Ltd Power Grid Corporation of	ranbaxy	14.76	271	0	22.84	798	0
India Ltd. Punjab National Bank	reliance industries	9.56	151	0	18.25	299	0
Ranbaxy Laboratories Ltd	reliance	1.39	15	0	2.00	41	0
Reliance Industries Ltd.	infrastructure sesa	2.23	35	0	2.95	83	0
Reliance Infrastructure	goa	35.06	444	0	59.70	591	0
Ltd. Sesa Goa Ltd.	sbi	0.54	15	0	0.61	8	0
State Bank of India	sun	40.45	440	1	72.47	647	0
Sun Pharmaceuticals Industries	pharmaceuticals tcs	27.69	204	0	31.48	187	0
		4.54	72	0	9.85	94	0
	tata motors	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE II. CORRELATION VALUES AND T STATISTICS FOR THE DIFFERENT COMPANIES TSTATISTICS(IN BRACKET), *

INDICATES STATISTICAL SIGNIFICANCE AT 0.05 (ONE TAIL).

Company	Correlation I		Correlation II		Correlation III		Correlation IV		Correlation V	
Ambuja Cements	0.13	(1.46)	0.11	(1.23)	0.46	(5.66)*	-0.14	(-1.46)	0.08	(0.85)
Ltd. Asian Paints	0.31	(3.63)*	-0.05	(-0.53)	0.24	(2.68)*	-0.23	(-2.49)*	0.16	(1.81)*
Ltd. Bajaj Auto	0.24	(2.72)*	0.02	(0.18)	0.47	(5.9)*	-0.28	(-3.08)*	0.28	(3.16)*
Ltd	0.42	(5.05)*	0.25	(2.86)*	0.47	(5.82)*	-0.27	(-2.93)*	0.25	(2.88)*
Bank of Baroda	0.28	(3.21)*	-0.03	(-0.34)	0.51	(6.43)*	-0.25	(-2.76)*	0.12	(1.32)
Bharat Heavy Electricals Ltd.	0.43	(5.17)*	-0.03	(-0.33)	0.68	(10.24)*	-0.35	(-3.93)*	0.32	(3.77)*
Bharat Petroleum Corporation	0.08	(0.85)	-0.04	(-0.42)	0.43	(5.09)*	-0.16	(-1.71)*	0.15	(1.65)
Ltd. Bharti Airtel Ltd.	0.00	(0.01)	0.12	(1.3)	0.04	(0.46)	-0.12	(-1.32)	-0.03	(-0.34)
Cairn India	0.13	(1.42)	0.09	(1.01)	0.67	(9.82)*	-0.26	(-2.84)*	0.18	(1.97)*
Ltd. Cipla	0.01	(0.11)	-0.03	(-0.35)	0.07	(0.8)	0.06	(0.66)	0.23	(2.6)*
Ltd.	0.05	(0.55)	0.17	(1.85)*	0.41	(4.93)*	-0.17	(-1.18)	0.21	(2.33)*
Coal India Ltd.	0.33	(3.83)*	0.06	(0.64)	0.16	(1.82)*	-0.10	(-1.08)	0.16	(1.82)*
Dr. Reddy's Laboratories	-0.09	(-0.99)	0.01	(0.1)	0.10	(1.15)	-0.02	(-0.18)	0.21	(2.41)*
Ltd. Gail (India) Ltd.	-0.12	(-1.33)	0.00	(-0.01)	0.54	(6.97)*	-0.31	(-3.54)*	0.27	(3.11)*
Grasim Industries	0.10	(1.12)	0.00	(-0.01)	0.21	(2.35)*	-0.19	(-2.13)*	0.13	(1.44)
Ltd. HCL	0.24	(2.69)*	0.10	(1.06)	0.03	(0.3)	-0.01	(-0.11)	-0.01	(-0.14)
Technologies Ltd.	0.27	(3.09)*	0.20	(2.19)*	0.27	(3.04)*	-0.28	(-3.03)*	0.10	(1.09)
HDFC Bank Ltd.	0.27	(3.01)*	0.35	(4.01)*	0.81	(14.88)*	-0.24	(-2.65)*	0.30	(3.45)*
Hero Motorcorp Ltd.	0.36	(4.19)*	0.00	(-0.05)	0.42	(5.02)*	-0.05	(-0.51)	0.38	(4.55)*
Hindalco Industries	0.54	(6.94)*	-0.01	(-0.15)	0.79	(14.07)*	-0.33	(-3.76)*	0.38	(4.52)*
Ltd. Hindustan	0.21	(2.34)*	0.11	(1.17)	0.33	(3.84)*	-0.28	(-3.09)*	-0.06	(-0.65)
Unilever Ltd. ICICI	0.23	(2.58)*	0.06	(0.7)	0.39	(4.6)*	-0.09	(-1.04)	0.26	(2.98)*
bank	-0.10	(-1.16)	-0.11	(-1.16)	0.04	(0.43)	-0.28	(-2.42)*	0.04	(0.42)
Infosys	0.26	(2.91)*	-0.09	(-1)	0.41	(4.99)*	-0.18	(-1.99)*	0.13	(1.47)
IDFC	0.11	(1.2)	0.23	(2.63)*	0.20	(2.3)*	-0.18	(-1.84)*	0.14	(1.5)
Ltd. ITC	0.24	(2.7)*	0.06	(0.71)	0.58	(7.86)*	-0.28	(-3.11)*	0.28	(3.25)*
Ltd.	0.33	(3.9)*	0.09	(0.95)	0.45	(5.59)*	-0.13	(-1.37)	0.08	(0.83)
Jaiprakash Associates	0.16	(1.73)*	0.12	(1.38)	0.38	(4.52)*	-0.21	(-2.25)*	-0.02	(-0.2)
Ltd. Jindal Steel Power	0.33	(3.79)*	0.15	-1.67	0.55	(7.16)*	-0.24	(-2.66)*	0.11	(1.22)
Ltd. Kotak Mahindra	0.06	(0.61)	0.04	(0.44)	0.11	(1.17)	0.00	(-0.03)	-0.20	(-2.22)
Bank Lupin Ltd.	-0.57	(-7.63)	0.04	(0.41)	0.83	(16.32)*	-0.10	(-0.9)	0.42	(4.96)*
Maruti Suzuki India	0.34	(3.94)*	0.14	(1.52)	0.68	(10.12)*	-0.59	(-5.49)*	0.33	(2.88)*
Ltd. NTPC Ltd.	0.05	(0.59)	0.38	(4.42)*	0.53	(6.82)*	-0.14	(-1.14)	0.03	(0.26)
Oil and Natural Gas Corporation	0.09	(0.98)	0.14	(1.59)	0.18	(2.07)*	-0.16	(-1.34)	0.15	(1.55)
Ltd Power Grid Corporation of	-0.10	(-1.12)	0.06	(0.66)	0.41	(5.01)*	-0.13	(-1.16)	0.25	(2.79)*
India Ltd. Punjab National Bank	0.55	(7.15)*	-0.04	(-0.41)	0.37	(4.4)*	0.01	(0.1)	0.34	(3.66)*
Ranbaxy Laboratories	-0.03	(-0.36)	0.05	(0.54)	0.29	(3.36)*	0.13	(0.89)	0.07	(0.77)
Ltd Reliance Industries	-0.11	(-1.25)	0.50	(6.19)*	0.53	(6.79)*	-0.33	(-3.69)*	0.30	(3.35)*
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

IV. CONCLUSION AND FUTURE WORK

The results show that Twitter sentiment is related to stock returns and tweet volume is also related to trading volume. Our results support the fact that when a stock is much talked about in Twitter the trading volume goes high. Incorporating information obtained from Twitter into financial models for predicting return, volatility and liquidity can be an area of further research. Here in this paper only Twitter as a social media source is considered. Further studies can include information from other social media like Facebook, LinkedIn etc.

Biographies

Muktamala Chakrabarti- Muktamala is currently a Doctoral student in the Management Information Systems group at Indian Institute of Management Calcutta (IIMC), India. Prior to joining IIMC, she completed her Bachelor of Electrical Engineering from Jadavpur University, India and worked at Tata Consultancy Services. Her research
interests include text analytics with applications in financial modeling.

Asim K. Pal- Asim K. Pal is a senior Professor at Indian Institute of Management Calcutta with the Management Information Systems area. He is a PhD in Computer Science from Calcutta University, India. He has published in several peer reviewed conferences and journals. His research interests include Business Intelligence, Data Mining, and Information Security.

Ashok Banerjee- Ashok Banerjee is currently the Dean (New Initiatives and External relations) of Indian Institute of Management Calcutta. He is a senior Professor with the Finance Group. He has publications in various peer reviewed journals and conferences.

REFERENCES

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31-56. Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*, *59*(3), 1259-1294.

Asur, S., & Huberman, B. A. (2010, August). Predicting the future with social media. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on* (Vol. 1, pp. 492-499). IEEE.

Baker, H. K., & Nofsinger, J. R. (2002). Psychological biases of investors. *Financial Services Review*, *11*(2), 97-116.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market.*Journal of Computational Science*, *2*(1), 1-8. Chamlertwat, W., Bhattarakosol, P., Rungkasiri, T., & Haruechaiyasak, C. (2012). Discovering Consumer Insight from Twitter via Sentiment Analysis. *J. UCS*, *18*(8), 973-992.

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*, *53*(9), 1375-1388.

Das, S., Martínez-Jerez, A., & Tufano, P. (2005). eInformation: A clinical study of investor discussion and sentiment. *Financial Management*, *34*(3), 103-137.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of business*, 34-105.

Gilbert, E., & Karahalios, K. (2010, May). Widespread Worry and the Stock Market. In *ICWSM* (pp. 59-65).

Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In*Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177). ACM.

Leinweber, D., and Sisk J.(2011) Relating news analytics to stock returns, in G. Mitra and L. Mitra (Eds.),*The handbook of news analytics in finance*, 596.

Lucey, B. M., & Dowling, M. (2005). The Role of Feelings in Investor Decision-Making. *Journal of economic surveys*, *19*(2), 211-237.

Mitra, L., & Mitra, G. (2011). Applications of news analytics in finance: A review. *The Handbook of News Analytics in Finance*, *596*, 1.

Mitra G., Dan diBartolomeo, Banerjee A., Yu X. (2011). Automated Analysis of News to Compute Market Sentiment: Its Impact on Liquidity and Trading. The Future of Computer Trading in Financial Markets - Foresight Driver Review – DR 8.

Refenes, A. P. (1994). *Neural networks in the capital markets*. John Wiley & Sons, Inc.

Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Transactions on Information Systems (TOIS)*, *27*(2), 12.

Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., & Deng, X. (2013, August). Exploiting Topic based Twitter Sentiment for Stock Prediction. In *ACL (2)* (pp. 24-29).

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

Tumarkin, R., & Whitelaw, R. F. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal*, 41-51.

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through twitter "I hope it is not as bad as I fear". *Procedia-Social and Behavioral Sciences*, *26*, 55-62