Chapter 8
The 'Return' and 'Volatility' of Sentiments: An Attempt to Quantify the Behaviour of the Markets?

Khurshid Ahmad

8.1 Introduction

Sentiment analysis is now becoming an established tool for the analysis of financial and commodity markets. The roots of this subject lie in the earlier work (c. 1950s) on content analysis on the one hand and on the other in work on bounded rationality and 'herd behaviour' by Herbert Simon and Daniel Kahnemann. Information extraction and corpus linguistics have been used in extracting the distribution of the so-called affect words and their collocates.

We begin this paper with a mini tutorial on the two key metaphorical terms used in finance studies, especially in the study of changes of prices and that of the value of the indices of a market: *return* and *volatility*. We briefly describe some related work in sentiment analysis. This is followed by a description of a corpus-based study of the variation in the frequency of positive and negative words, as defined in the *Harvard Dictionary of Affect*. An afterword concludes the paper.

8.2 Metaphors of 'Return' and of 'Volatility'

The literature on financial economics that is closely related to the analysis of market sentiment frequently refers to two key terms: *return* and *volatility*. These terms have retained much of their original meaning that is, 'return' broadly refers to the 'act of coming back', as established upon the entry of this word in the English language in the fourteenth century or thereabouts. This word has been adapted, or to put loosely has been used as a metaphor for coming back as in the definition: 'Pecuniary value resulting to one from the exercise of some trade or occupation; gain, profit, or income, in relation to the means by which it is produced'. In financial economics a return, or 'price change quantity' is defined as the logarithm of the ratio of the

K. Ahmad (✉)
School of Computer Science and Statistics, Trinity College, Dublin 2, Ireland
e-mail: kahmad@scss.tcd.ie

K. Ahmad (ed.), *Affective Computing and Sentiment Analysis*, Text, Speech and Language Technology 45, DOI 10.1007/978-94-007-1757-2_8,

current and past price (or an index of a stock exchange or any other aggregated index, like Standards & Poor, Financial Times-Stock Exchange Index):

Let p_t be the price today and p_{t-1} the price yesterday, so the return r_t is defined as:

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right)$$

The word 'volatility' is much more graphic as it started its journey into the English language from Latin in the seventeenth century: 'The quality, state, or condition of being volatile, in various senses' and the metaphorical use of this word includes references to 'tendency to lightness, levity, or flightiness; lack of steadiness or seriousness'. Benoit Mandelbrot [14] has argued that the rapid rate of change in prices (the *flightiness* in the change) can and should be studied and not eliminated – 'large changes [in prices] tend to be followed by large changes – of either sign – and small changes tend to be followed by small changes'. The term *volatility clustering* is attributed to such clustered changes in prices. Mandelbrot's paper drew upon the behaviour of commodity prices (cotton, wool and so on), but volatility clustering is now used in for almost the whole range of financial instruments (see [17] for an excellent and statistically well-grounded, yet readable, account of this subject).

There are different kinds of measures of volatility, a commonly used version is called *realized* or *historical volatility*. Volatility (v) of a stock price or the value of an index is defined over a trading period n and is the standard deviation of the past returns ($\ln(\frac{p_t}{p_{t-1}})$):

$$v = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(r_{t-k} - \bar{r})}$$

where $\bar{r}$ is the average value of past returns in the period n. Econometricians have observed that it is perhaps easier to conduct a statistical analysis of changes in market prices as there is little or no correlation between consecutive price changes; there is a significant correlation in the prices. It has been argued by Robert Engle [9], the 1993 co-winner of the Nobel prize in economics, that '[a]s time goes by, we get more information on these future events and re-value the asset. So at a basic level, financial price volatility is due to the arrival of new information. Volatility clustering is simply clustering of information arrivals. The fact that this is common to so many assets is simply a statement that news is typically clustered in time.' (1993:330).

The term *news arrivals or information arrivals* is defined rather differently in the literature in finance. This standard practice in financial economics is either to use daily counts of news stories as a proxy for information arrival or, in simulations, a random number generator is used for generating the number of news arrivals [5]. This method has been refined to count only those stories that comprise a given (set of) keyword(s) [6, 7], and more recently Tetlock [18] has used affect words in the

General Inquirer lexicon [16] to correlate market movements with change in the frequency of affect words. More of this later.

Let us look at the concepts of return and volatility in the context of *information arrivals* by looking at the recurrence of news items containing the same keywords and, more importantly, on the recurrence of the same metaphorical terms. Consider the use of the term *credit crunch*: According to Wikipedia (2008) the term is defined as ‘a sudden reduction in the availability of loans (or “credit”) or a sudden increase in the cost of obtaining a loan from the banks.’ This term has been around for over 30 years or more, but let us look at the usage of this term since 1981: The *New York Times* archives were searched for the compound term ‘credit crunch’ and over 400 articles, published between 1981–2008, were found that contained the term. There may have been other articles which the NYT authorities did not or could not include, but the NYT is considered as an authoritative and usually un-biased source of US and international political and economic reporting. The number of stories containing the term appear to have a 5 year cycle, except in the last decade where ‘credit crunch’ only appeared in large number in 2007; in 2008 there were 54 stories compared to whole of 2007 where there were 77 stories in all. My projected value of the number of stories in 2008 is 156, based on the current average of about 13 a month (Table 8.1).

Let us look at how the ‘returns’ of the number of stories in Table 8.1 and compute the annual historical volatility purely on the basis of the changes in the numbers of stories in *New York Times*. It has been argued that volatility, computed using return of prices or index values, increases during crises periods, say during the 1929 US Great Depression, the period leading upto the resignation of the US President Nixon in 1974, and the days after the 2001 9/11 attacks (see [17, pp. 191–93]).

We have plotted the number of stories per year, the consecutive year returns, and the volatility for a reporting (‘trading’?) period of 5 years. There is a much greater variation in the returns (based on the current and past numbers of news stories containing the term ‘credit crunch’) when compared to the fluctuations recorded in the actual number of stories. One quantification of such a fluctuation is volatility or the standard deviation of past returns. The volatility increased every 5 years until 2000 – indicating an *increasing* sense of ‘crises’. Volatility decreases during calmer periods. In our extremely simple illustration, we note that the volatility decreased during 2001–2005 – a period of massive growth partly fuelled by ‘easy credit’ – and lastly

Table 8.1 The number of stories per year comprising the term *credit crunch* that appeared in (or are in the archive of) *New York Times*

Year	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	Decade Total
# Stories	19	6	4	4	8	5	3	4	6	59	118
Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	
# Stories	93	38	24	4	10	3	2	47	3	4	228
Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	
# Stories	10	4	1	2	1	2	225	616	168	59	1088

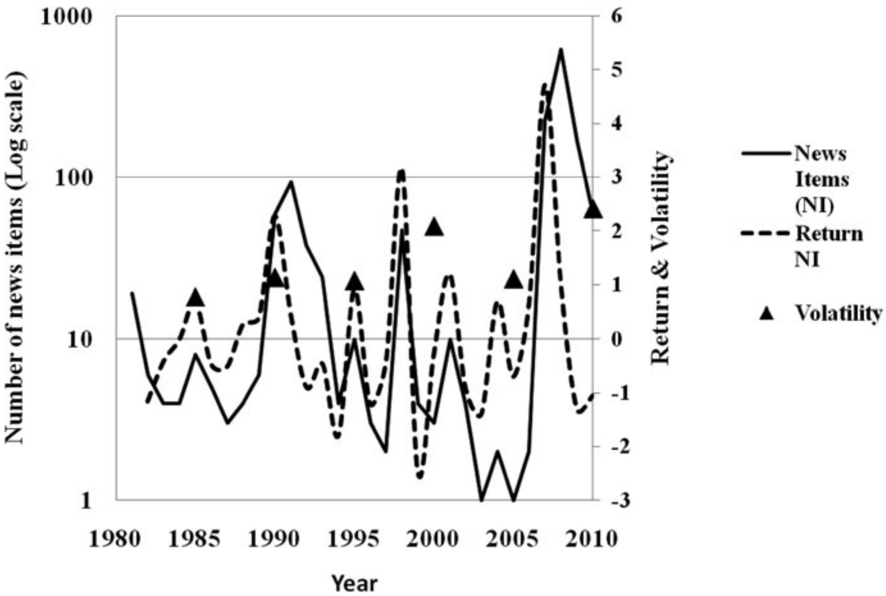

Fig. 8.1 The number of stories are on the left vertical stories, and the percentage change in returns and volatility are on the right

for the 5 years (2006–2010) the volatility has increased dramatically incorporating the period of the very frequent use of the term 'credit crunch' (Figure 8.1):

The above figure shows that a single statistic, volatility, can be used in the quantification of rapid changes. The questions is this: will this statistic throw any light on the changes in market sentiment, based on methods in sentiment analysis that use the frequency of positive and negative affect words as a measure of sentiment?

8.3 The Roots of Computational Sentiment Analysis

Sentiment is defined as 'an opinion or view as to what is right or agreeable' and political scientists and economists have used this word as a technical term. When sentiments are expressed through the faculty of language, we tend to use certain literal and metaphorical words to convey what we believe to be right or agreeable. There are a number of learned papers and reviews in computational sentiment analysis that are available ([11] and references therein).

One of the pioneers of political theory and communications in the early twentieth century, Harold Lasswell [12], has used sentiment to convey the idea of an attitude permeated by feeling rather than the undirected feeling itself. (Adam Smith's original text on economics was entitled *A Theory of Moral Sentiments*.) Namenwirth and Laswell [15] looked at the Republican and Democratic party platforms in two periods 1844–1864 and 1944–1964 to see how the parties were converging and how language was used to express the change. Laswell created a dictionary of affect words (*hope*, *fear*, and so on) and used the frequency counts of these and other words to quantify the convergence.

This approach to analysing contents of political and economic documents – called content analysis – was given considerable fillip in the 1950s and 1960s by Philip Stone of Harvard University who created the so-called General Inquirer System [10, 16] and a large digitised dictionary – the *GI* Dictionary comprising over 8,500 words carefully selected using a criterion developed by the psychologist Charles Osgood including positive/negative words, words to express strength and weakness, and words to describe activity and arousal (Stone’s dictionary includes a number of entries used by Harold Laswell; these entries are thus labelled).

Recently, the digitised Harvard *Dictionary of Affect* has been used to ‘measure’ sentiment in the financial markets. Tetlock [18] has analysed a commentary column in the *Wall Street Journal* using the GI Dictionary and correlated the frequency of affect words with trading volumes of shares in the New York Stock Exchange: He concludes that ‘high media pessimism predicts downward pressure on market prices followed by a reversion to fundamentals, and unusually high or low pessimism predicts high market trading volume.’ This is amongst the first reported study in financial sentiment analysis that is rooted strongly in econometric analysis (especially through the use of auto-regressive models in the framework of conditional heteroskedasticity) that has analysed the contents of the news in conjunction with the study of information arrival (see also [13]). Tetlock’s selection of comment or opinion in a newspaper, classified as imaginative writing rather than the informative news reportage, may raise some methodological questions in text analysis about whether or not opinions can in themselves comprise a representative sample of texts that has been used for analysis (see, for example [2, 4]).

8.4 A Corpus-Based Study of Sentiments, Terminology and Ontology Over Time

We report on some work recently carried on compiling a representative or random sample of texts, in a given domain, that can be used for analysing sentiments. Once the a corpus is compiled we then extract terminology that is used in the domain automatically. Then significant collocates of the candidate domain terminology are used in the construction of a candidate ontology (see [1] for details). In my previous work I have avoided using pre-compiled dictionaries of affect and used the so-called ‘local grammar’ constructs for extracting patterns that were ‘sentiment-laden’ [3] – an approach that has allowed us to look for sentiment in texts in typologically diverse languages like English (and Urdu), Chinese and Arabic. For the purposes of comparison with other work in financial sentiment analysis, I have used the Harvard *Dictionary of Affect*: I have computed returns and volatility of affect in a corpus drawn from a representative newspaper website. My hypothesis is this: Can the computation of the volatility of affect, found in news paper reportage and editorials, help in quantifying (financial and economic) risk, much in the same manner as risk computations based on prices and values of index help in quantifying risk?

8.4.1 Corpus Preparation and Composition

The design of the corpus was motivated by the state of Irish economy during the period of 1995–2005; the first 5 years were the so-called Celtic Tiger boom eras (1996–2000) and the next 5 year period comprised the dot.com bubble, September 2001 attacks and the consequent down turn and the lead upto the introduction of the Euro. The authoritative and influential *Irish Times* that has been published since the 1850s and has a digital archive going back to 1859. One of my student (Nicholas Daly) used a text-retrieval robot to search and retrieve all items (news reports, editorials, or op-ed columns) based on the robot user: We chose the period (1995–2005) and gave the robot three keywords: *Ireland/Irish* and *Economy*. The corpus comprises 2.6 million words distributed over 4,075 news reportage and editorial items (Table 8.2):

The size of the year, viewed on an annual basis, appears to be comparable (Mean = 407, Standard Deviation = 51.522): only in 2 years both the number of stories and the verbiage was above one-standard deviation above the mean (1996 and 2001), and the number of stories in 2005 were just one s.d. above the mean (1.04).

8.4.2 Candidate Terminology and Ontology

We found that 'sentiment' in itself was a keyword and analysis of its statistically significant collocates showed that despite the boom in the late 90s the focus of *Irish Times* content was on more negative aspects, but the next 5 years show the establishment of a whole terminology nucleating around 'sentiment' (Table 8.3):

Table 8.2 Distribution of stories in our *Irish times corpus*

Year	No. of stories	No. of words	Year	No. of stories	No. of words
1996	296	1,65,937	2001	562	3,60,026
1997	395	2,59,748	2002	367	2,56,613
1998	465	2,96,531	2003	377	2,50,415
1999	447	2,95,873	2004	377	2,50,376
2000	462	3,06,063	2005	327	2,34,101
Total	2,065	13,24,152		2,010	13,51,531

Table 8.3 Compound words with 'sentiment' as a head word – a comparison over 5 year periods

1995–2000	2001–2005
▼ • sentiment	▼ • sentiment
▼ • investor_sentiment	▶ • business_sentiment
• factors_affecting_investor_sentiment	▼ • consumer_sentiment
▼ • market_sentiment	▶ • consumer_sentiment_survey
▶ • bond_market_sentiment	▶ • irish_consumer_sentiment
▶ • negative_sentiment	▶ • investor_sentiment
▶ • poor_sentiment	▼ • sentiment_index
	▶ • sentiment_index_climbed
	▶ • sentiment_index_produced
	▶ • sentiment_index_rose
	▶ • sentiment_index_surpassed

The above analysis was carried out using the computation of significant collocates following Frank Smadja (1993) and the assumption here was that if the word sentiment is to the left of a another word, excluding the so-called closed class words, then sentiment is the headword. The output was processed using the ontology system Protègè.

8.4.3 Historical Volatility in Our Corpus

The 2.6 million word corpus was analysed by computing the frequency of affect words in Harvard *Dictionary of Affect* (H-DoA) that were present in the texts in the corpus. The frequency was normalised for the length of the individual texts. The H-DoA comprises a large number of categories as mentioned above: we have used only two categories, *Positive* and *Negative* affect word categories that respectively have 1,916 and 2,292 words. For each news item on a given day, the frequency of all words that were labelled *Positive* and *Negative* in the H-DoA was computed. The frequency counts were aggregated on a monthly basis and returns computed. The standard deviation of the returns on annual basis was calculated and we then had *volatility* of 'positive' sentiments and that of the 'negative' sentiments.

The first thing to notice about our results is that the 'return' (change in frequency) shows much greater fluctuation in value than the frequency itself; this confirms the findings in econometrics in the context of prices and the change in prices (see, for example, 2003). This is true of both the negative and positive word frequency time-series, despite the preponderance of positive words over the negatives (Figs. 8.2 and 8.3).

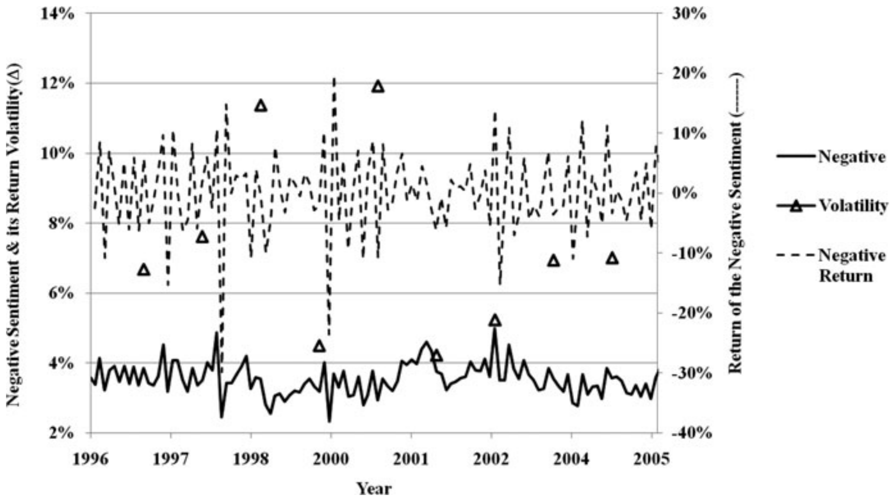

Fig. 8.2 Changes in the frequency (*full line*) of negative affect terms in our *Irish Times Corpus* (displayed monthly for 1996–2006). The returns are shown in *dashed line* (and values on the vertical axis on the right hand. The historical volatility is indicated by *solid triangles* and values are on the left-hand vertical axis

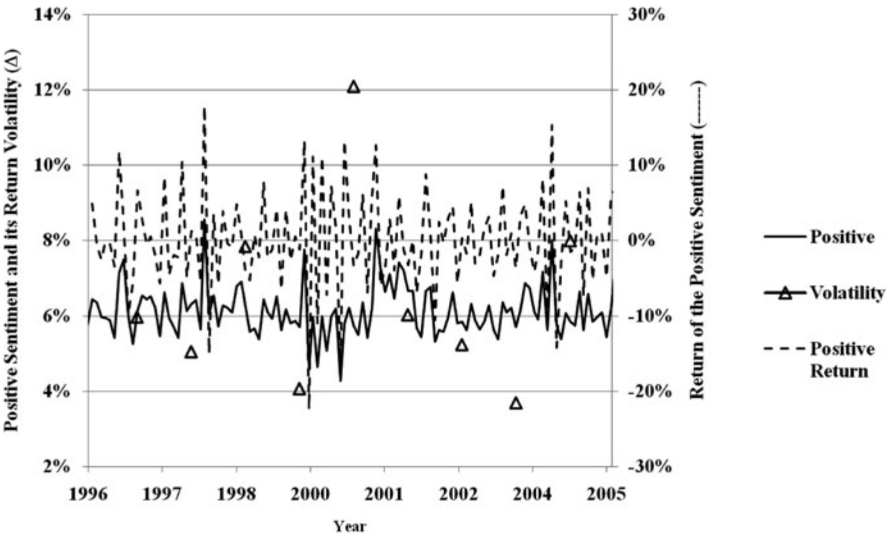

Fig. 8.3 Changes in the frequency (*full line*) of positive affect terms in our *Irish Times Corpus* (displayed monthly for 1996–2006). The returns are shown in *dashed line* (and values on the vertical axis on the right hand. The historical volatility is indicated by *solid triangles* and values are on the left-hand vertical axis

The changes in historical volatility computed over the 10 year period shows interesting results: in 1998 and 2000 the negative series had a higher than 'normal' volatility (one standard deviation above the norm) and in the two intervening years the volatility was below the norm (1999 and 2001). The positive affect series has below the norm volatility in 1999 and 2003 and much higher volatility in 2000 (2 standard deviations) (see Table 8.4).

Finally, we show the variation in the Irish Stock Exchange Index of 100 top companies listed on the Exchange (ISEQ 100). We have had access to the values of the Index on a daily basis and we have used the value at the end of the month of

Table 8.4 Volatility changes in our two time series

	Volatility		Volatility	
Year	Negative	Std. dev.	Positive	Std. dev.
1996	0.064		0.057	
1997	0.075		0.050	
1998	0.112	1.7	0.078	
1999	0.034	−1.2	0.038	−1.1
2000	0.111	1.6	0.113	2.2
2001	0.036	−1.1	0.060	
2002	0.054		0.052	
2003	0.054		0.037	−1.1
2004	0.070		0.079	
2005	0.058		0.059	

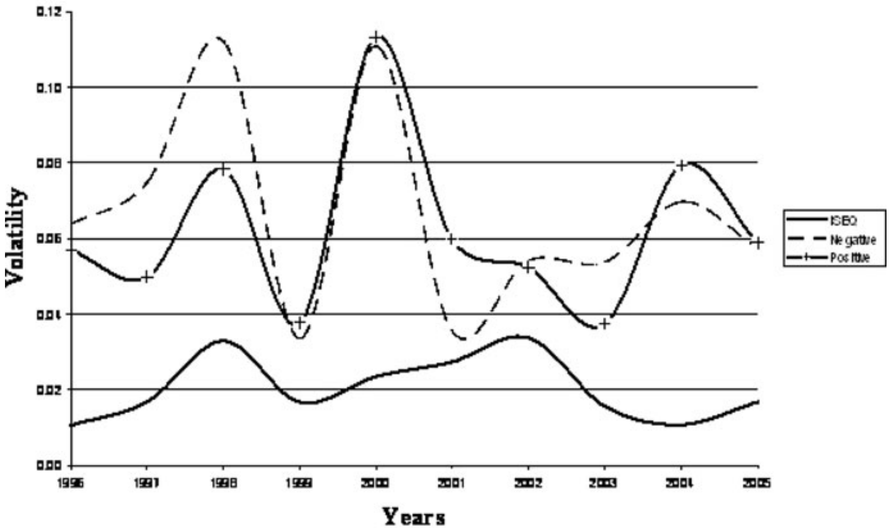

Fig. 8.4 Changes in the historical volatility in the affect series and in the ISEQ Index

each year as the ISEQ Index value and then computed returns and volatility for the period 1996–2005.

The volatility in ISEQ is smaller in comparison with that in the negative and positive affect series: this may be an artefact of computation as is the considerable variation in the volatility series of affect when compared to ISEQ (see Fig. 8.4).

8.5 Afterword

The above results give us some sense of how to find sentiment words and quantify the changes in sentiment. It is perhaps too early to read the runes: whether we can use the volatility of affect times series to compute (financial) risk. But the study looks promising. We are looking at the auto-correlation in the various time series and computing other econometric metrices to quantify changes in sentiment.

In a related study, myself and my colleagues are looking at the effect of the use of different dictionaries of affect on the measurement of sentiments, including that of the GI Dictionary. We hope to use the system for analysing reports about emerging markets and specific financial instruments (shares, derivatives, bonds) and commodities: we intend to go beyond the professional media (newspapers, company documents, stock exchange reports) and include social media (blogs, e-mails and contrarian reports). It is through the social media that the contagion affecting the stock markets spreads. This project is undertaken jointly with Trinity Business School and the Irish Stock Exchange.

A sentiment analysis based on the indirect evidence of social and professional media is only one part of the overall picture. The Trinity Sentiment Analysis Group, a multi-disciplinary group including computer scientists, linguists and economists, has launched a sentiment survey for Irish institutional and individual investors. This survey was originally developed by Robert Shiller of Yale University International Centre of Finance; we have launched this Survey in collaboration with Yale.[1] The work of the Trinity Sentiment Group is ambitious and is focussed on engendering an openness and transparency in the workings of the vitally important financial sector. We are endeavouring to bring together and synthesise inputs from the professional media, the social media, data from the stock markets, and views of the stakeholders in a common framework. This is a long term program of work which we have just begun.

Acknowledgments Dr Ann Devitt worked with me on this project and we have looked at different dictionaries of affect [8]. She had written the program that computes affect word frequency.

References

1. Ahmad, K. 2007. Artificial ontologies and real thoughts: Populating the semantic web? (Invited Keynote Speech). In *Lecture notes on artificial intelligence series (4733) – Proceedings of 10th Annual Congress of Italian Ass. of Art. Intelligence*, eds. R. Basili and M.T. Pazienza, 3–23. Berlin: Springer.
2. Ahmad, K. 2007. Being in text and text in being: Notes on representative texts. In *Incorporating Corpora: The linguist and the translator*, eds. G. Anderman and M. Rogers, 60–94. Clevedon, England: Multilingual Matters.
3. Almas, Y., and K. Ahmad. 2007. A note on extracting 'sentiments' in financial news in English, Arabic & Urdu, The Second Workshop on Computation, al Approaches to Arabic Scriptbased Languages, Linguistic Society of America 2007 Linguistic Institute, Stanford University, Stanford, California., Linguistic Society of America, pp. 1–12.
4. Althaus, S., J. Edy, P. Phalen. 2001. Using substitutes for full-text news stories in content analysis: Which text is best? *American Journal of Political Science* 45(3):707–723.
5. Bauwens, L., W.B. Omrane, and P. Giot. 2005. News announcements, market activity and volatility in the Euro/Dollar foreign exchange market. *Journal of International Money and Finance* 24:1108–1125.
6. Chang, Y., and S.J. Taylor. 2003. Information arrivals and intraday exchange rate volatility. *Journal of International Financial Markets, Institutions and Money* 13:85–112.
7. DeGennaro, R., and R. Shrieves. 1997. Public information releases, private information arrival and volatility in the foreign exchange market. *Journal of Empirical Finance* 4:295–315.
8. Devitt, A., and K. Ahmad. 2007. Cohesion-based sentiment polarity identification in financial news. The 45th annual meeting of the Association for Computational Linguistics (ACL 2007), Prague, Czech Republic, June 23–30, 2007, Stroudsburg, PA: Association of Computational Linguistics, 984–991.
9. Engle III, R. 2003. *Econometric models and financial practice.* http://nobelprize.org/nobel_prizes/economics/laureates/2003/engle-lecture.html http://en.wikipedia.org/wiki/Credit_crunch. Accessed 17th April 2008.
10. Kelly, E., and P. Stone.1975. *Computer recognition of English word senses*. Amsterdam: North-Holland Linguistic Series.

[1]https://www.cs.tcd.ie/Khurshid.Ahmad/SurveySite/

11. Kennedy, A., and Inkpen, D. 2006. Sentiment classification using contextual valence shifters. *Computational Intelligence* 22(2):117–125.
12. Lasswell, H.D. 1948. Power and personality. London: Chapman and Hall.
13. Lidén, E.R. 2006. Stock recommendations in swedish printed media: Leading or misleading? *The European Journal of Finance* 12(8):731–748.
14. Mandelbrot, B. 1963. The variation of certain speculative prices. *Journal of Business* 36: 394–419.
15. Namenwirth, Z., and H.D. Lasswell. 1970. *The changing language of American values: A computer study of selected party platforms*. Beverly Hills CA: Sage Publications.
16. Stone, P.J., D.C. Dunphy, M.S. Smith, and D.M. Ogilvie, et al. 1966. *The General inquirer: A computer approach to content analysis*. Boston:The MIT Press. (For downloading the GI Lexicon, please go to http://www.wjh.harvard.edu/~inquirer/. Accessed 17 April 2008.
17. Taylor, S.J. 2005. *Asset price dynamics, volatility, and prediction*. Princeton and Oxford: Princeton University Press.
18. Tetlock, P.C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62(3):1139–1168.