

Swiss Finance Institute  
Research Paper Series N°13-01

# The sentiment of the Fed

Michel FUKSA  
AGH University of Science and Technology

Didier Sornette  
ETH Zurich and Swiss Finance Institute

swiss:finance:institute



National Centre of Competence in Research  
Financial Valuation and Risk Management

Established at the initiative of the Swiss Bankers' Association, the Swiss Finance Institute is a private foundation funded by the Swiss banks and Swiss Stock Exchange. It merges 3 existing foundations: the International Center FAME, the Swiss Banking School and the Stiftung "Banking and Finance" in Zurich. With its university partners, the Swiss Finance Institute pursues the objective of forming a competence center in banking and finance commensurate to the importance of the Swiss financial center. It will be active in research, doctoral training and executive education while also proposing activities fostering interactions between academia and the industry. The Swiss Finance Institute supports and promotes promising research projects in selected subject areas. It develops its activity in complete symbiosis with the NCCR FinRisk.

The National Centre of Competence in Research "Financial Valuation and Risk Management" (FinRisk) was launched in 2001 by the Swiss National Science Foundation (SNSF). FinRisk constitutes an academic forum that fosters cutting-edge finance research, education of highly qualified finance specialists at the doctoral level and knowledge transfer between finance academics and practitioners. It is managed from the University of Zurich and includes various academic institutions from Geneva, Lausanne, Lugano, St.Gallen and Zurich. For more information see [www.nccr-finrisk.ch](http://www.nccr-finrisk.ch).

This paper can be downloaded without charge from the Swiss Finance Institute Research Paper Series hosted on the Social Science Research Network electronic library at:

<http://ssrn.com/abstract=2191937>

# **The sentiment of the Fed**

Michel Fuksa

AGH University of Science and Technology, Cracow Poland (mwebmail@gmail.com)

Didier Sornette

ETH Zurich, Department of Management, Technology and Economics, Zurich, Switzerland  
(dsornette@ethz.ch)

## **Abstract**

Sentiment analysis is applied to the Federal Reserve's documents relating to the Federal Open Market Committee (FOMC). A new method of rating the sentiment of texts is developed through the creation of a 100 000 word sentiment dictionary, and is used to estimate the sentiment of the Beige Book and FOMC statements and minutes. Results show that, under Greenspan's tenure, the sentiment of the Beige Book anticipated the sentiment of the FOMC minutes while, during the Bernanke era, the sentiment of the Beige Book anticipates the sentiment of the FOMC statement. This finding suggests that the Beige Book, usually made public three weeks before a FOMC meeting, may contain information useful in anticipating policy decisions.

**JEL classification:** C53, E58, G00, G17

**Keywords:** economics, Federal Reserve, Federal Open Market Committee, Beige Book, word sentiment dictionary, prediction

# 1-Introduction

Sentiment analysis is a tool for examining the emotional content of texts. It has found many uses in academia and industry, see Weber (1983).

In finance, Tetlock (2007, 2008) was the first to apply sentiment analysis to the analysis of textual financial information. By comparing the sentiment-content of the Wall Street Journal's "Abreast of the Market" column with the response of the market, Tetlock (2007) found that pessimism in that particular piece of text predicted downward market movement. Moreover, Tetlock (2008) found that the negativity of stock-related stories could effectively predict market movement.

Recently, Twitter sentiment analysis has become popular because the nature and use of this medium requires the transmission of a maximum of information in a limited amount of space. The popularity of this analysis for Twitter is also a function of the high frequency of this medium and its short response time to events. O'Connor et al. (2010) for example applied sentiment analysis to Twitter messages, consumer surveys and political polls, and found that Twitter sentiment reliably measured opinions.

Similarly Turney (2002), Zhuang et al. (2006) and Schuller and Knaup (2011) used sentiment-analysis methods to analyze movie reviews. Dodds and Danforth (2010) proposed a method for examining the emotions of different types of text documents based on the ANEW study (Bradley and Lang 1999) and found significant results in a variety of different texts.

The idea behind this type of analysis is that the choice of words used in expressing an opinion or making a statement is a function of the mood or state of mind of the writer. Moreover, a writer may compose a text to attempt to influence his audience through a careful selection and matching of words.

Sentiment is in a way inseparable from content: words carry not only sense but emotion and two texts can potentially contain similar information-content but have distinct impact because of the choice of words.

In this article, sentiment analysis is applied to documents relating to the Federal Reserve's Federal Open Market Committee. In particular sentiment analysis is used to examine whether the mood of FOMC statements and minutes is related to the mood of the Beige Book. The rest of the article is organized as follows. Section 2 reviews the literature on lexical analysis of Federal Reserve documents, which allows us to position the present work in its context. Section 3 describes the used databases and the corpus of analyzed texts and the existing methods to rate sentiments. Section 4 introduces our new tool for sentiment analysis, and tests it on the Moby valence dictionary. Section 5 reviews the main characteristics of the Federal Open Market Committee (FOMC) communications, the obtained sentiments and presents a preliminary analysis of its characteristics. Section 6 presents quantitative metrics of the FOMC text sentiments and analyses the properties of its time series, such as auto- and cross-correlations. Section 7 concludes.

## **2-Previous Lexical analysis of Federal Reserve texts and position of the present work**

There have been several attempts to glean insight from Federal Reserve documents and relate it to monetary policy or movements in financial asset prices. These studies have focused on two documents: the Beige Book and the minutes of the Federal Open Market Committee (FOMC). The Beige Book is known more formally as the Summary of Commentary on Current Economic Conditions, and is a report published by the United States Federal Reserve Board eight times a year. A short review of the literature on the subject is now presented.

One of the earliest analyses of the kind was performed by Payne (2001), who examined the message of the Beige Book's leading paragraph for each Federal Reserve district for insights into monetary policy action. The authors found that the Beige Book was a fair predictor of economic conditions and helps predict the change in coincident and leading indicators. The measure constructed on the basis of lexical information also appears to slightly anticipate the movements in the Federal Funds rate.

Balke and Petersen (2002) manually constructed an index to evaluate the content of the Beige Book. The Beige Book was found to explain GDP growth over and above Blue Chip forecasts available in real time. The authors suggest that the Blue Book contains information about current and next quarter GDP growth above and beyond what is available in other sources.

Zavodny and Ginther (2005) examined the influence of the release of the Beige Book on a variety of financial prices. They quantified the national summary of the Beige Book in a manner similar to Balke and Petersen (2002) and then correlated the so obtained quantitative index with financial asset prices. The results showed that the Beige Book does have some impact on changes in intermediate and long term interest rates, but has no impact on equities. It was found that the Beige Book appears to have no explanatory power over future monetary policy.

Sadique et al. (2010) use the General Inquirer (Stone et al. (1966)) to analyze the tone of the Beige Book. It is found that the tone of the Beige Book is a function of economic condition: the Beige Book has a positive tone in economic upswings and negative in downswings. The authors find that the Beige Book has some influence on short-term interest rates and on movements in the SP500 index.

Armesto et al. (2009) reexamined the evidence on the ability of the Beige Book to predict economic aggregates. Using DICTION software (Hart, 2000) and measuring the optimism and pessimism of the text, they found that the Beige Book can explain 15 to 18% of GDP growth and that optimistic Beige Book language predicts positive GDP growth and inversely. It was found that aggregate employment could be explained with similar power.

Boukus and Rosenberg (2006) applied latent semantic analysis to the text of the FOMC minutes. Their analysis showed that the minutes reflect complex economic information and that they can be useful as a source of information in predicting economic activity. They discover that several themes are related to current market indicators, especially three-month yields and GDP growth.

In the end, the analysis of the Beige Book and FOMC minutes appears inconclusive. There is evidence of some influence on financial asset prices and there is some evidence for the predictive power of the content of these documents regarding economic variables. The strength of this evidence is in question since different studies use different methods of rating documents and often contradict themselves.

Several features distinguish the current work from previous literature. First, the present study focuses on the relation between documents issued prior and post FOMC meetings. To date, the literature has focused on the relation of FOMC-related documents to the outside world as represented by economic indicators and financial asset prices. In this study, the interrelation between FOMC-related documents is explored. It is hoped that new light will be shed into the nature of communications relating to monetary policy.

In this article, sentiment analysis is chosen as the tool for analyzing Federal Reserve documents pertaining to the Federal Open Market Committee. It is believed that this approach will yield new

insight into relations between FOMC documents as compared with lexical analysis. The motivation for this expectation is that sentiment analysis quantifies lexical complexity of documents and that numerical analysis may make it easier to spot interesting relationships in text data.

The third aspect of this study is the use of a large sentiment dictionary compared to the literature. Previous studies evaluate documents either manually, or using two well recognized software systems: the General Inquirer (Stone et al. (1966)) and DICTION (Hart (2000)). While these systems represent the gold standard in lexical analysis, their limiting feature are their relatively small dictionaries (around 10 000 words each). This study develops a 100 000 word sentiment dictionary patterned on the ANEW study (Bradley and Lang, (1999)) to score documents.

### **3-Sentiment corpora and databases**

In order to analyze the sentiment in a text or corpus of texts, a method of rating sentiment is first required. The standard method is to define a sentiment dictionary composed of a list of words and the corresponding sentiment-related ratings. There exist several starting points for this type of analysis.

The gold standard resource for this type of task is the General Inquirer (GI) dataset of Stone et al. (1966). Another possibility is the ANEW dataset project of Bradley and Lang (1999). The GI dataset is a classic resource and contains a list of close to 10000 words tagged with markers representing select categories. Among categories of interest for this study are those classifying the words into positive and negative sentiments. A disadvantage of the GI dataset is that it does not assign a numerical rating or scale to sentiment.

The ANEW study (Bradley and Lang (1999)) on the other hand, developed a numerical rating scale to grade the positivity and negativity of words. The words were rated on a scale from 1 to 9, and



measured three dimensions: affective valence (good - bad), arousal and dominance and control.

Unfortunately, the ANEW vocabulary is of limited scope (1040 words) and because of that it may not be able to authoritatively rate random documents.

In finance, Armesto (2008) used the DICTION software package (Hart (2000)) to analyze text. This software is not suitable for the type of analysis proposed in this study, however, since it does not publish the word lists it uses.

Dodds et al. (2011) and Kloumann et al. (2012) used a crowd-sourcing technique to assess the sentiment of a selected number of words and have made public the results. Their work, patterned and validated against the ANEW study contains "happiness" ratings for 10,222 unique words. The vocabulary was released publicly as the labMT 1.0 dataset (Dodds et al. (2011)). The vocabulary contains the most frequent words found in the Twitter, The Google Books Project (English), The New York Times, and Music lyrics corpora. The vocabulary consists of a list of words assigned with a "happiness" score from 1 to 9, as well as a standard deviation of that score, a remnant of the automated word rating process. The 9-point rating scale, originating in the ANEW study, gives low scores (closer to 1) to negatively charged words and high scores (closer to 9) to positive words. On this scale, a perfectly negative word would rate 1, and a perfectly positive word would rate 9. In the dataset, no such extremes are present, but some words come close.

The labMT 1.0 vocabulary provides numerical ratings of sentiment. It also is representative of the modern English language because it is composed of words from popular English language sources. For these reasons, this dataset was chosen as the source sentiment dictionary for the study. All further computing in this study was performed with the help of the Natural Language Toolkit (Bird, 2009).

## **4-A new tool for sentiment analysis**

### ***4.1 Description of the new 100'000 word sentiment dictionary***

The problem with many of the afore-mentioned word lists is that they are limited in scope and include only a small number of rated words. In fact, it would be impossible to manually assign an affective rating to every word in the English language. Fortunately, there exist methods for extrapolating sentiment scores to a large number of words starting from a seed list.

This is accomplished by boosting, a process where sentiment ratings are assigned to new words based on a seed dictionary. A well recognized method is to find new words synonymous to words present in the seed dictionary and to assign them the same sentiment rating.

The method of boosting-by-synonyms was chosen for this study. To do this, a thesaurus was used to find synonyms to words in the labMT seed vocabulary and a procedure was elaborated to assign ratings to the synonyms. The thesaurus used in the boosting procedure is the freely available Moby thesaurus by Ward (2002) containing over 100 000 words. The Moby thesaurus is a set of lists of related words. Following these relations, a network was constructed consisting of thesaurus words as vertices and their relations as edges. For example, if two words are related, they would be linked in a simple graph consisting of two vertices, each vertex representing a word, and an edge linking the two vertices. This procedure led to the creation of a single graph encoding the Moby thesaurus into 102,682 vertices representing words and 2,519,426 edges representing relations between words.

The first step in rating the sentiment of words encoded in this network structure was to identify common words in the labMT and Moby vocabularies and to assign an initial sentiment rating to vertices representing these common words. The second step was to use the network structure of the thesaurus to propagate these initial ratings to the remaining unrated words, resulting in a large

sentiment corpus. The procedure for propagating the sentiment rating values was to iterate over the unrated vertices in the thesaurus network and, for each unrated vertex, to compute a rating as an average of the ratings of the neighboring vertices. Experimentation has shown that this procedure leads to stable results after only 100 iterations over the whole network. Starting from 10,222 rated sentiment-words in the labMT vocabulary, the boosting procedure just described produced a set of 102,682 sentiment-rated words on the scale from 1 to 9. The rating scale in the new dictionary is the same as that in the seed dictionary, that is, low sentiment words have ratings towards 1 while high sentiment words have ratings towards 9. In the remainder of the article, the new sentiment vocabulary will be referred to as the Moby sentiment thesaurus.

#### ***4.2 Testing the Moby valence dictionary***

A first look at the Moby dataset reveals that most words have no strong positive or negative connotations and that most words, as shown in Figure 1, have a rating close to the average of 5.4. This means that many words are close to sentiment-neutral and it may be difficult to cleanly classify certain words as positive or negative, especially if they fall at the boundary rating of 5.4. On the other hand, the slight skewness of the distribution indicates that the vocabulary of the English language is optimistic, which implies that any text analyzed through this method will have a small positive bias. The negative skew evidenced through the long tail on the left side of the distribution suggests that a proportion of the vocabulary has very negative meaning. This implies that texts containing these words will be strongly affected and will tend to show pessimistic scores.

In order to test whether the boosting procedure produced correct results, the newly produced vocabulary was tested against the General Inquirer (Stone et al. (1966)) valence categories and against the Hu and Liu (2004) list of positive and negative sentiment English words.

The General Inquirer (GI) has two valence categories: words with positive outlook and words with

negative outlook marked with 4 tags ('Positiv', 'Pstiv', 'Negativ' and 'Ngtv'). The test consisted in finding common words between the two vocabularies (a total of 3590 common words were found) and verifying whether GI words with positive outlook scored higher on the Moby scale than GI words tagged with negative outlook. A similar procedure was repeated using the Hu and Liu (2004) dataset, which provides two word lists, one for positive words and one for negative words.

It is found that the Moby and the GI dictionaries rate words in a similar way with very little misclassification (about 7% total). In short, GI words marked as positive have a Moby rating greater than 5.4 and negatively tagged GI words have scores less than 5.4. The Hu and Liu word lists are classified in the same manner: there is a 7.5% misclassification rate (414 words out of a total of 5459 words).

A cursory look at misclassified words reveals two sources of error. First, some words are effectively misclassified. Second, it appears that the meaning of the word plays a role in its rating. Some words may have multiple meanings, some positive, some negative. It appears that the Moby valence dictionary rates some "meanings" in a different way than either the GI or Hu and Liu word lists. A possible cause of error is the fact that the seed labMT dictionary was rated by the general public while the comparative lists were rated by professionals. Thus, the Moby thesaurus tends to reflect the "popular" meaning of words.

In all, the small misclassification rate suggests that the crowd-sourcing method used for developing the original labMT vocabulary as well as the thesaurus based boosting technique produced relatively unbiased results and can be used with confidence to rate the sentiment of texts.

Overall, the synthetic Moby dataset can be seen as performing a good job in classifying words into the positive and negative valence categories. The value of 5.4 is chosen as the rating discriminating

between positive and negative sentiment General Inquirer words. This value is extrapolated to classify the entire Moby corpus into two large valence categories: Moby words with a rating of 5.4 or less are considered as having negative sentiment, whereas Moby words with a rating greater than 5.4 are considered as having positive sentiment. The leptokurtic distribution of word sentiment means that the most words have a sentiment value close to the mean sentiment score and that there are relatively few words with extreme sentiment loadings. This implies on the other hand that the sentiment of a text will depend mainly on “mild” words. The high kurtosis of the sentiment distribution also means that performing an analysis excluding these average-sentiment words will lead to the exclusion of the majority of the words in the sentiment dictionary and will render the rating of texts difficult if not impossible. For example, there are 2820 words with a sentiment rating of less than 4, and only 789 words with a sentiment greater than 7. Using a coarser text-rating procedure based on this partition will not create a reliable scoring system since most of these words are not related to economics.

The new Moby vocabulary displays some advantages compared to other databases. Whereas the Harvard General Inquirer dataset only identifies words as either of positive or negative valence, the Moby dataset uses a continuous rating on the scale of 1 to 9. Moreover, the Moby corpus is simpler in construction, broader and appears to be as reliable as its older counterpart. While the General Inquirer dictionary classifies words according to various categories above and beyond positive and negative sentiment, it is a little difficult to use since the interpretation of analyzed texts may pose a problem. A single rating like the one provided by the Moby corpus makes the interpretation of text sentiment analysis easier.

The General Inquirer should not be sidelined however for the simple reason that it is a gold standard in content analysis as it provides the qualitative information not available elsewhere. Its drawback is the limited number of words it contains and the rather awkward disambiguation of

meanings of certain keywords. Ideally, a one-to-one correspondence between WordNet (Fellbaum (1998)) and General Inquirer would allow for a detailed analysis of almost any text, but such an option is not available as of today.

## **5- Financial Open Market Committee (FOMC) communications and sentiment data**

### ***5.1 A short overview of FOMC communications***

The Federal Open Market Committee (FOMC) has a long and complex history of communicating with the public (see Danker, Luecke (2005) and Bernanke (2007)). Since 1981, the FOMC meets 8 times per year (previously, twelve times per year). Surrounding each meeting, a set of documents is released, including the Beige Book, the FOMC statement and the FOMC minutes. These documents are not issued synchronously at the end of the meeting, but their publication is spread out in time around the date of the meeting. Over the years, the changing philosophy affecting these documents has affected their consistency and content.

From the year 2000 onwards, a statement is released following each FOMC meeting. The statement is composed of a set of five paragraphs succinctly presenting the decisions taken at the meeting as well as the reasons behind them. The first two paragraphs present a short discussion of economic conditions and economic outlook while the third and fourth paragraphs discuss the federal funds rate and the balance sheet. Finally, votes in favor of the statement are counted and any dissent is recorded. The FOMC statements are continually evolving and have recently started to include participant projections for the federal funds rate and discussions of long-term policy goals (see Guha (2012)).

The FOMC minutes are a summary of the reasoning behind policy decisions and are published three weeks after each regularly scheduled meeting (prior to 2004, the publication delay was over

30 days). The minutes consist of views and arguments about the economic outlook and policy that can not be included in statements. The minutes are a record of "all policy-related topics" (Dankert and Luecke (2005)) discussed by the participants and reflect the information available at meeting time.

Starting from 1983, the Beige Book is produced 8 times per year and is released 2 weeks prior to each FOMC meeting. The Beige Book is a structured document composed of a summary of economic conditions at the national level and a set of reports for each of the twelve Federal Reserve districts.

## **5.2 Sentiments of FOMC communications**

Sentiment analysis using newly developed Moby vocabulary was applied to the documents of the Federal Reserve's Federal Open Market Committee (FOMC) presented in Table 1. The data for the study is composed of documents from the end of the Greenspan tenure and over the Bernanke tenure at the Federal Reserve, that is, from the January 2000 until February 2012.

Due to the evolution of Federal Reserve communications, the FOMC document record has been inconsistent over time. Since the release of the first FOMC statement in 1994, not every FOMC meeting has released a statement or a record of minutes. For example, in 1994, a complete document set was released at only 5 meetings, whereas 1996 and 1997 featured only one statement per year. It is only in 2000 that scheduled FOMC meetings started to consistently release a statement and a record of minutes. On the other hand, the Beige Book is available eight times per year over this period. Also, as mentioned in the previous paragraph, these documents have continued to evolve over this time span.

Given these considerations, the text corpus is chosen as a set of 98 meeting triples (Beige Book,

FOMC statement and FOMC minutes) released synchronously over the period from January 2000 to February 2012. The dataset spans the end of the Greenspan tenure (up to and including the FOMC meeting on January 31 2006) and the Bernanke tenure to date (starting from the FOMC meeting on March 27-28 2006). The dataset thus includes 49 meetings from the Greenspan tenure and 49 meetings from the Bernanke tenure.

The analysis focuses on the estimation of sentiment for the three documents. The FOMC statements and minutes are analyzed separately. The Beige Book is composed of a national summary and 12 district summaries. Only the national summary is analyzed in this document.

### ***5.3 Preliminary analysis of FOMC communications***

A first look at the data has revealed that the three corpora under analysis (Beige Book, FOMC statements and minutes) have evolved over time. The size of the documents composing the corpora was measured by counting the number of tokens, or individual strings of characters, composing each text.

The results show that the length of all FOMC documents has grown over the study period. The average number of tokens in the Beige Book national summary was 1858 in the year 2000, while it was 2899 in 2011 - a 55% increase. Similarly, the length of the FOMC statements and minutes increased significantly: there was a 150% increase in the length of the FOMC statements and a 63% increase in the length of the minutes over the same period. On average, the Beige Book grows by 8.7 tokens per meeting, the FOMCS statement by 3.4 words, and the FOMCS minutes by 44.5 words between regularly scheduled FOMC meetings.

More significantly, there does appear to be a difference in communication style between the Greenspan and Bernanke tenures. Regressions illustrating a trend break between the two tenures



show this point clearly. To illustrate this point, document length was regressed on a linear trend variable and a dummy variable taking the value zero for the period of the Greenspan tenure (first 49 observations) and then taking the value of a linear trend for the Bernanke tenure (the last 49 observations).

The results show that the two Fed chairmen have two very different communication styles (see Table 2 and Figure 3). In accord with the pledge of more transparency (see Bernanke (2007)), the length of FOMC statements and minutes has increased more significantly under Bernanke than under Greenspan. The Bernanke Fed is more verbose and provides the public with more information than the Greenspan Fed. This suggests that the current FOMC, with its new style, may also present information differently than before.

The FOMC statement data shows a clear break at the beginning of 2006 when the new chairman took control of the Fed. Under the new leadership, the length of the FOMC statements increased significantly and kept on increasing until today. Both the trend and the dummy trend variables are highly significant in the trend regressions. Under Greenspan, the length of the statements increased by 1.5 tokens per meeting, while under Bernanke it is 5 tokens per meeting.

A similar story holds for the FOMC minutes. Under Greenspan, the length of the FOMC minutes appears to decrease slightly but the trend is not actually significant. There is however a very significant increase in the length of FOMC minutes during the Bernanke tenure, with the length of meeting minutes increasing by 120 tokens per meeting.

In the case of both the statements and minutes, the trend does not appear to abate. In any case, there is strong evidence for an increased communication intensity by the Federal Reserve under Bernanke.

Concerning the Beige Book, there does not appear to be any trend break at the time of change of tenure. This is motivated by the fact that the Beige Book is a Federal Reserve system document mandated by law and is produced by a bureaucratic process insensible to leadership changes. The Beige Book remains a stable document over time and is unperturbed by subjective or personal styles of the Board of Governors.

The other regressions presented in Table 2 examine if the length of the documents changed with the advent of the financial crisis in mid-2008. There does appear to have been a change in the amount of communications from the Federal Reserve after the start of the crisis. The regressions concur however that, even before the beginning of the crisis, there was an increase in Fed documents. The advent of the crisis then significantly augmented the amount of information put out by the Fed.

This evidence points to a general change in the information-content of documents released by the Fed under chairman Bernanke.

## **6-Computation and analysis of text sentiment**

### ***6.1 Computing text sentiment***

Starting with the two basic ingredients, a valence dictionary and a text document, the sentiment of the text can be computed using a simple algorithm. In this paper, two measures of sentiment were computed using these resources: a sentiment score measuring the average mood of a text, and a tone score, measuring the positive or negative attitude of a text.

For both metrics, the first step in the procedure is to determine the frequency of occurrence of each word in the text. The sentiment score  $V_{text}$  of a text is computed by calculating a weighted average

of individual word valences taken from the valence dictionary and the frequency of word occurrence. This procedure is presented by the following formula, where the sums are taken over all the words in the text that are also in the Moby dictionary,  $f(w)$  is the frequency of a word and  $r(w)$  is its rating:

$$V_{text} = \frac{\sum_w f(w) * r(w)}{\sum_w f(w)} \quad (1)$$

The second metric, called 'tone'  $T_{text}$ , tells if a text exhibits a positive or negative attitude by contrasting the number of positive and negative words it contains. For the purposes of this measure, a word is considered as "positive" if its valence is above the mode of the Moby sentiment distribution (Moby sentiment value above 5.5), and negative otherwise:

$$T_{text} = \frac{\text{number of positive words} - \text{number of negative words}}{\text{number of positive words} + \text{number of negative words}} \quad (2)$$

## 6.2 Sentiment time series

Sentiment was computed for the documents presented in Table 1 for each FOMC meeting over the period January 2000 to February 2012. A sentiment score and a tone score were produced for every FOMC minute, statement and Beige Book national summary released over this time.

First, an ANOVA analysis of the sentiment before and after the financial crisis (Table 3) shows that the sentiment expressed in Federal Reserve documents has decreased during the crisis period. All three documents examined in the study show the same behavior, which suggests that the texts express insight about the economic situation.

Figures 4 and 5 present the graphs of the computed sentiment for the FOMCS statements and minutes as well as for the Beige Book. Figure 4 presents sentiment ratings obtained through the

Moby sentiment dictionary and Figure 5 shows the tone of the documents.

It can be noted that sentiment ratings using the Moby dictionary show less and oscillate close to the mean sentiment rating of 5.5. The tone measure produces more variability.

### **6.3 Trend**

There is evidence of a downward trend in sentiment and tone in the FOMC minutes and statements data over the period covered by the study. There does not appear to be a break in the trend at the time of change of leadership at the Fed.

There is not evidence of a trend in the sentiment or tone of the Beige Book series and it appears to be a consistent measure of economic sentiment.

The downward trend in the two mood measures indicates either worsening economic conditions over time, or an attempt at balancing the mood of the documents relating the economic situation. To complete this point, the downward trend in sentiment for both the minutes and statements shows an average decrease from 5.6 to 5.4 points on the 1 to 9 sentiment scale. Similarly, the average tone decreases from 0.2 to close to 0 over the period 2000-2012. This suggests that the documents have become more neutral over time, while they were more positive historically (a tone of 0 implies the equal presence of positive and negative words, and 5.4 is the average sentiment rating in the Moby valence dictionary).

### **6.4 Cross-correlations**

In order to check for any potential relations between the sentiment ratings of the Beige Book and the subsequent FOMC statements of minutes, a series of correlation tests and regressions were performed.

There appears to be a small but significant correlation between the sentiment of the Beige Book and the sentiment of both the statement and minutes (see Table 4). Specifically, the correlation is negative between the sentiment of the Beige Book and FOMC statements and it is positive between the Beige Book and FOMC minutes. Looking at the situation more closely, namely dividing the data into the Greenspan and Bernanke periods, there appears to have been a rather strong positive correlation between the Beige Book and minutes in the Greenspan era (correlation of 0.5), and there is a strong negative correlation between Beige Book sentiment and FOMC statement sentiment in the Bernanke tenure (correlation of -0.44).

The interpretation of the situation is complicated by the fact that the Beige Book is released on average about 3 weeks before a FOMC meeting takes place. It appears that it may be currently possible to predict the sentiment of the FOMC statement by examining the sentiment contained in the Beige Book. The Beige Book may thus be construed to constitute to some extent a leading indicator of the sentiment of the Federal Reserve.

Another intriguing aspect of these results is that at the time of change of leadership at the Federal Reserve, there appears to be a shift in the relation between Beige Book sentiment and FOMC meeting documents. During the Greenspan era, there appears to have been some relation between the Beige Book and the FOMC meeting minutes, currently however, there is some relation between the Beige Book and the FOMC statement.

Some of this may be explicable by the changes in communication strategy pursued by the Fed during the Bernanke tenure. The increased length of the FOMC statements may include more information about the policy of the Federal Reserve and this may in turn explain the power of the Beige Book in predicting this sentiment.

## **6.5 Auto-correlations**

Figure 6 shows that sentiment time series of Federal Reserve monetary policy documents (as displayed in Figure 4) are significantly positively auto-correlated. The partial autocorrelation functions display a pattern characteristic for an ARMA (1,0) model.

The graphs further show that the sentiment time series are cross-correlated with each other with some lag. This implies that the future behavior of one of the time series could be potentially predicted from the behavior of the other sentiment time series.

It is interesting to note that the Beige Book sentiment is negatively correlated in time with the sentiment of the FOMC statement but is positively cross-correlated with the sentiment of the FOMC minutes. It appears that the Federal Reserve is releasing statements containing information on policy direction while the minutes contain discussions about current economic conditions.

## **7-Conclusions**

Sentiment analysis is a method for extracting the emotional meaning of words from texts. In this article we applied sentiment analysis to the documents produced by the Federal Reserve in the setting of monetary policy.

It was found that there exist some unexpected regularities in the sentiment of the FOMC statements and the Beige Book. In essence, it was found that the sentiment of the Beige Book leads the sentiment of the FOMC statement by a period of about three weeks. Furthermore the sentiment of these two documents is significantly auto-correlated leading one to suspect that economic policy is not set randomly and that it may be possible to predict to some extent the economic sentiment of the Federal Reserve's FOMC. The potential for exploiting these findings for economic gain is still to be explored.

## References

- Armesto, Michelle T, Rubén Hernández-Murillo, Michael T Owyang, and Jeremy Piger. 2009. Measuring the Information Content of the Beige Book: A Mixed Data Sampling Approach. *Journal of Money, Credit and Banking* 41, no. 1: 35-55. <http://dx.doi.org/10.1111/j.1538-4616.2008.00186.x>.
- Balke, Nathan S., and D'Ann Petersen. 2002. How well does the Beige Book reflect economic activity? Evaluating qualitative information quantitatively. *Journal of Money, Credit, and Banking* 34:114–36.
- Bernanke, Ben S. (2007) Federal Reserve Communications, The Cato Institute 25th Annual Monetary Conference, Washington, D.C., November 14, 2007
- Bird, Steven, Klein, Ewan, Loper, Edward, Natural Language Processing with Python, Publisher: O'Reilly Media, June 2009
- Boukus, Ellyn, and Joshua V Rosenberg. 2006. The information content of FOMC minutes. *New York* 18, no. 1: 1-12.  
[http://www.newyorkfed.org/research/economists/rosenberg/Boukus\\_and\\_Rosenberg\\_072006.pdf](http://www.newyorkfed.org/research/economists/rosenberg/Boukus_and_Rosenberg_072006.pdf)
- Bradley, M.M., & Lang, P.J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. Technical report C-1, Gainesville, FL. The Center for Research in Psychophysiology, University of Florida.

Danker, Deborah J., Luecke, Matthew M. (2005) Background on FOMC Meeting Minutes, Federal Reserve Bulletin, Spring 2005

Dodds PS, Danforth CM (2010) Measuring the Happiness of Large-Scale Written Expression: Songs, Blogs, and Presidents, *Journal of Happiness Studies*, Volume 11, Number 4 (2010), 441-456

Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM (2011) Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. *PLoS ONE* 6(12): e26752. doi:10.1371/journal.pone.0026752

Fellbaum, Christiane (1998, ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Guha, K. (2012) Changes in FOMC communications, Federal Reserve Bank of New York, March 19, 2012

Hart, Roderick P., *DICTION 5.0: The Text-Analysis Program*, Scolari/Sage Publications, Thousand Oaks, CA, 2000.

Kloumann IM, Danforth CM, Harris KD, Bliss CA, Dodds PS (2012) Positivity of the English Language. *PLoS ONE* 7(1): e29484. doi:10.1371/journal.pone.0029484

Minqing Hu and Bing Liu. "Mining and Summarizing Customer Reviews." *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004)*, Aug 22-25, 2004, Seattle, Washington, USA



O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *Proceedings of the International AAAI Conference on Weblogs and Social Media* (Vol. 5, p. 122–129). AAAI Press.

Payne, David R., 2001, Anticipating Monetary Policy with the Federal Reserve's Beige Book: Re-specifying the Taylor Rule, *Business Economics* 21-29.

Sadique, Shibley, Francis Haeuck In, Madhu Veeraraghavan, and Paul Wachtel. "Soft Information and Economic Activity: Evidence from the Beige Book." SSRN eLibrary (August 31, 2010). [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1669724](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1669724).

Schuller, B., Knaup, T., (2011) Learning and Knowledge-Based Sentiment Analysis in Movie Review Key Excerpts, Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues, Lecture Notes in Computer Science, Springer Berlin / Heidelberg

Stone, Philip J; Dexter C. Dunphry; Marshall S. Smith; and Daniel M. Ogilvie. (1966) *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.

Tetlock, Paul, (2007) "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *Journal of Finance* 62, no. 3 (June 2007): 1139-1168.

Tetlock, Paul, Maytal Saar-Tsechansky, and Sofus Macskassy, (2008) "More Than Words: Quantifying Language to Measure Firms' Fundamentals." *Journal of Finance* 63 (June 2008): 1437-1467.

Turney, P.D. (2002) Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, pp. 417–424 (July 2002)

Ward, Grady, (2002), The Project Gutenberg Etext of Moby Thesaurus II by Grady Ward, Edition 1.0, <http://www.gutenberg.org/ebooks/3202>

Weber, Robert Philip, (1983) Measurement models for content analysis, *Quality & Quantity*, Volume 17, Number 2 (1983), 127-149

Zavodny, M., and D. K Ginther. “Does the Beige Book Move Financial Markets?” *Southern Economic Journal* (2005): 138–151.

Zhuang, L., Jing, F., Zhu, X.Y. (2006) Movie review mining and summarization. In: Proceedings of the 15th ACM International Conference on Information and Knowledge Management (CIKM 2006), pp. 43–50. ACM, New York

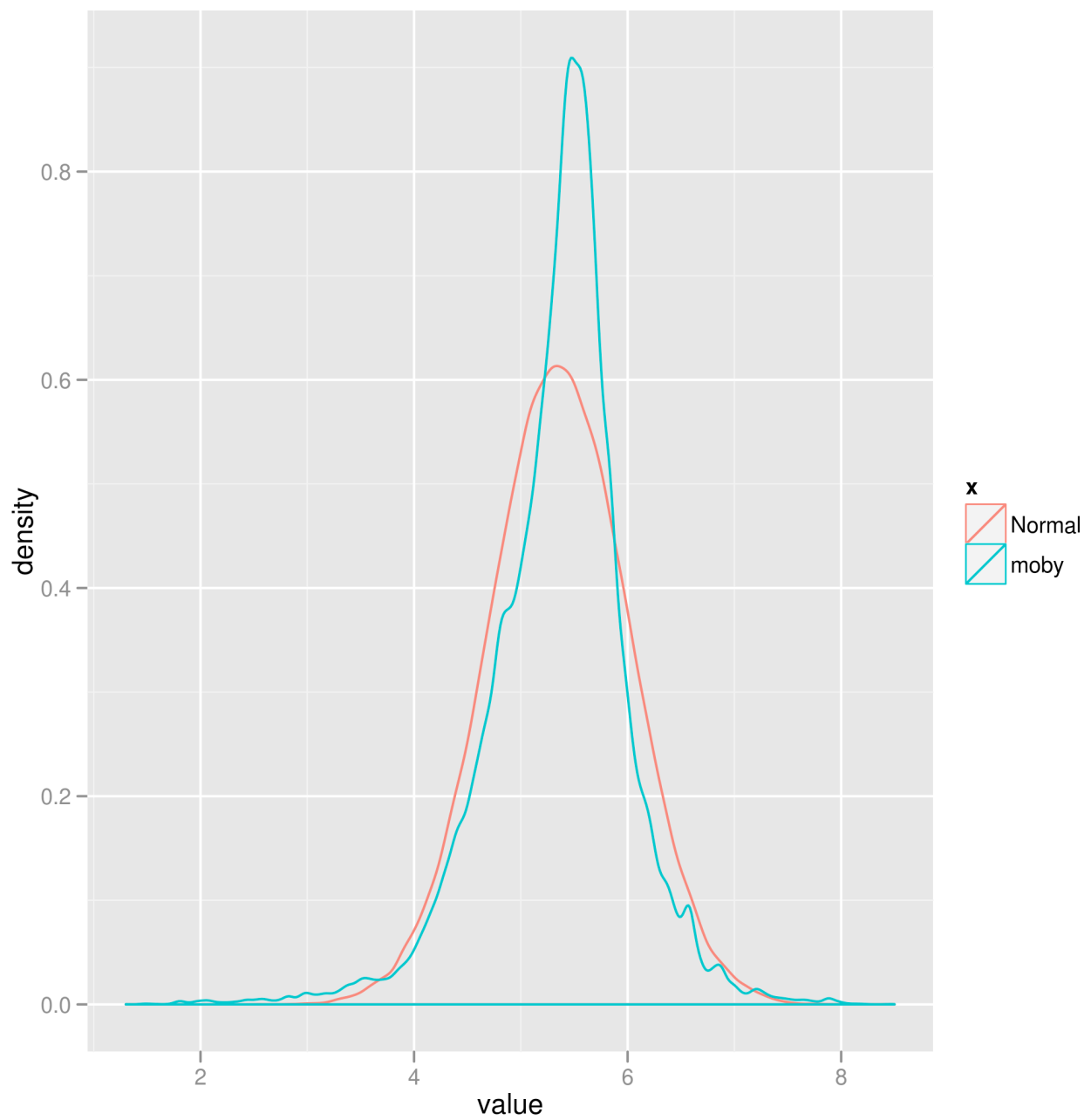


Figure 1. The distribution of word sentiment rating for the 102 682 words of the Moby sentiment dictionary contrasted with a Normal curve of the same mean (5.35) and standard deviation (0.64). The skewness of the sentiment distribution curve is -0.6 and its kurtosis 2.83. The horizontal axis shows Moby sentiment rating on the scale from 1 to 9, while the vertical axis shows word frequency.

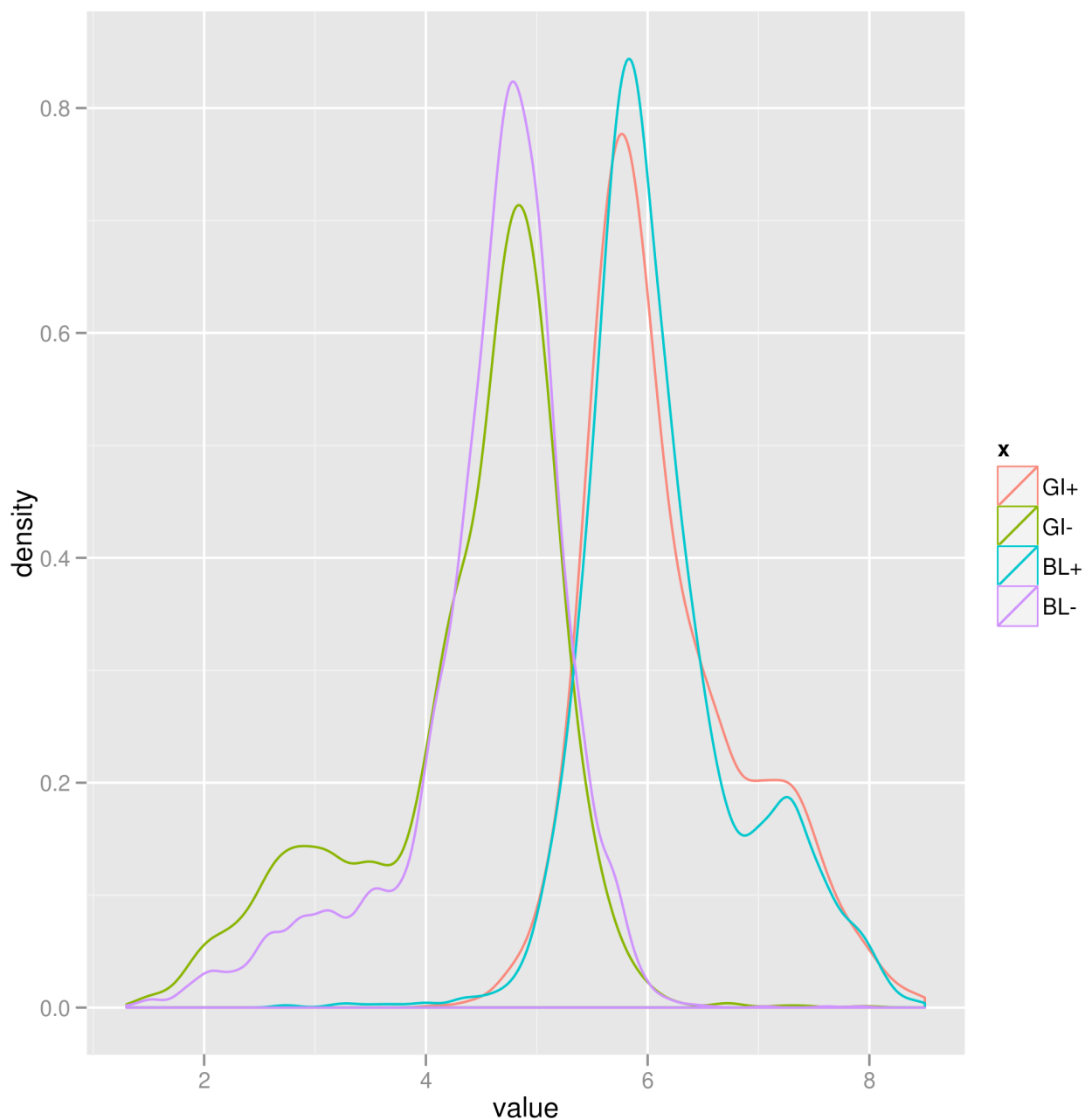


Figure 2. The figure illustrates the power of the Moby sentiment dictionary to discriminate between positive and negative words from the General Inquirer and Hu and Liu (2004) reference datasets. Positive words tend to have higher sentiment ratings than negative words. The horizontal axis shows Moby sentiment rating (on a scale from 1 to 9), while the labels GI+ and GI- stand for words classified respectively as positive and negative by the General Inquirer, and labels BL+ and BL- stand for words classified respectively as positive and negative by Hu and Liu (2004).

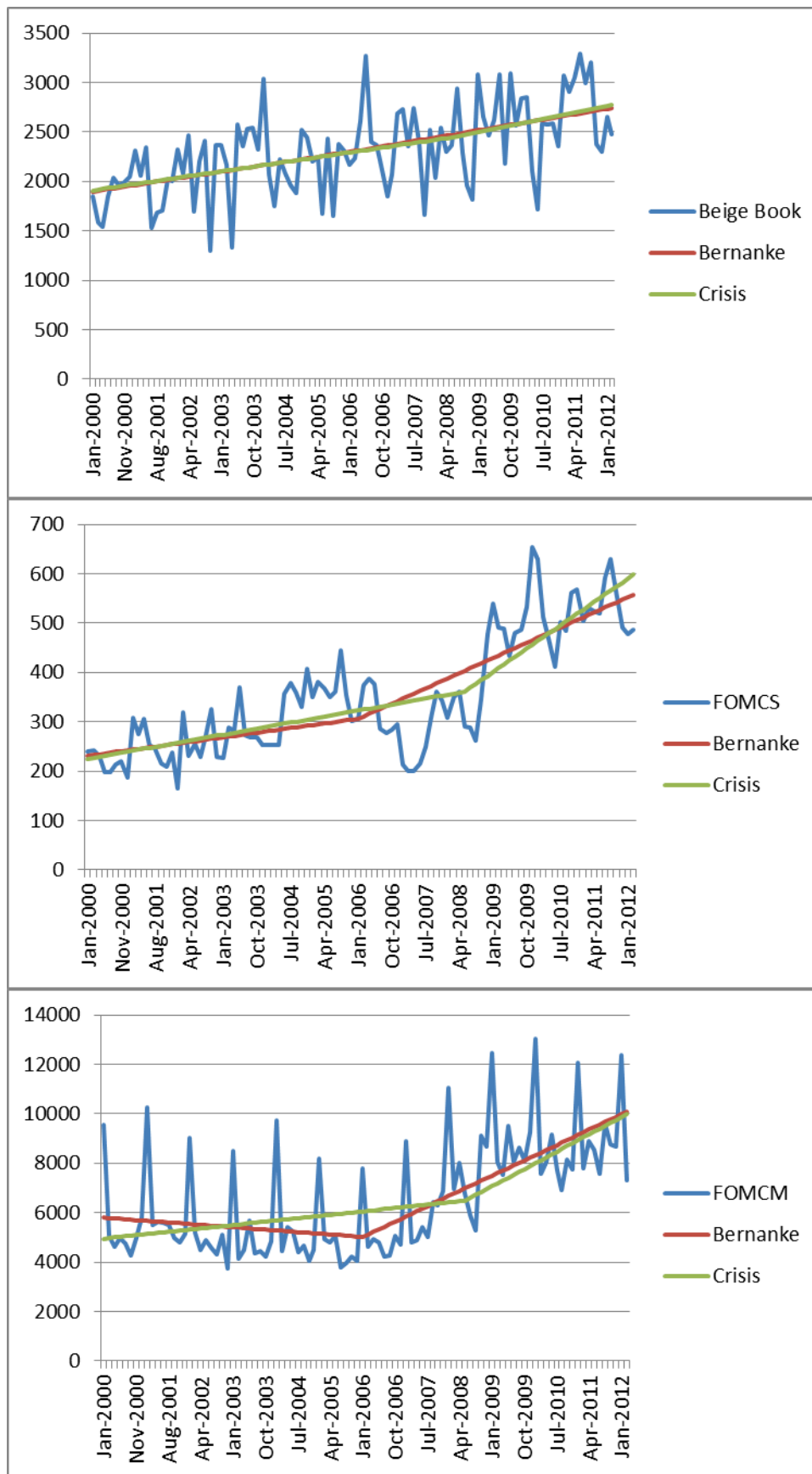


Figure 3. The graphs show data and related trend models for the size of FOMC related documents as presented in Table 2. The lines ‘Bernanke’ and ‘Crisis’ represent the regressions given in the

caption of Table 2.

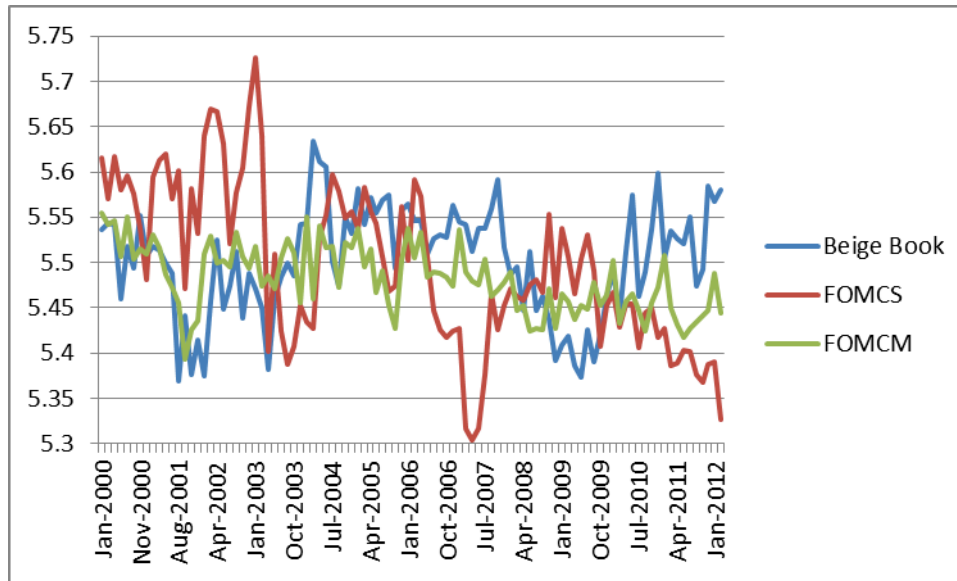


Figure 4. Sentiment ratings for the three major FOMC documents: Beige Book national summary, FOMC statement and FOMC minutes.

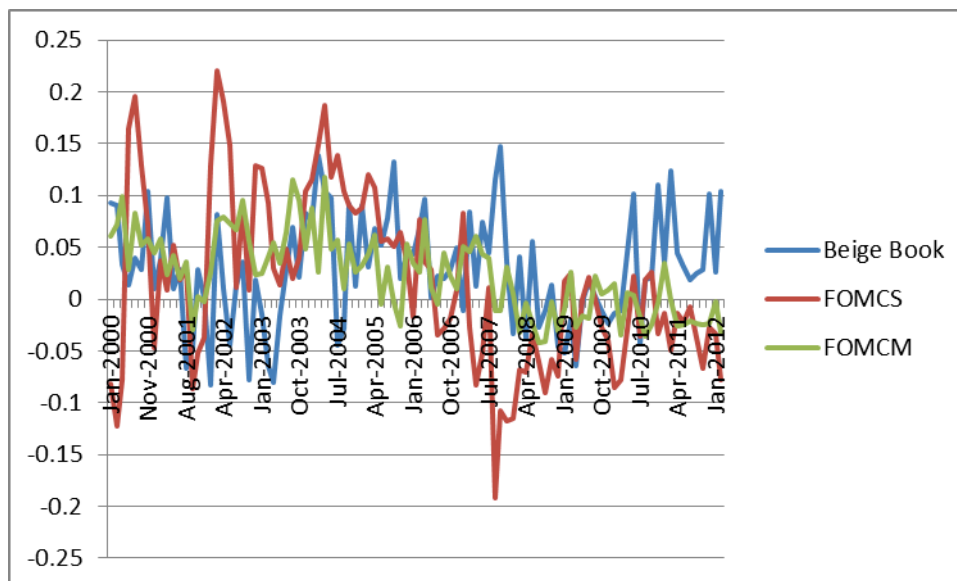


Figure 5. Tone ratings for the three major FOMC documents: Beige Book national summary, FOMC statement and FOMC minutes.

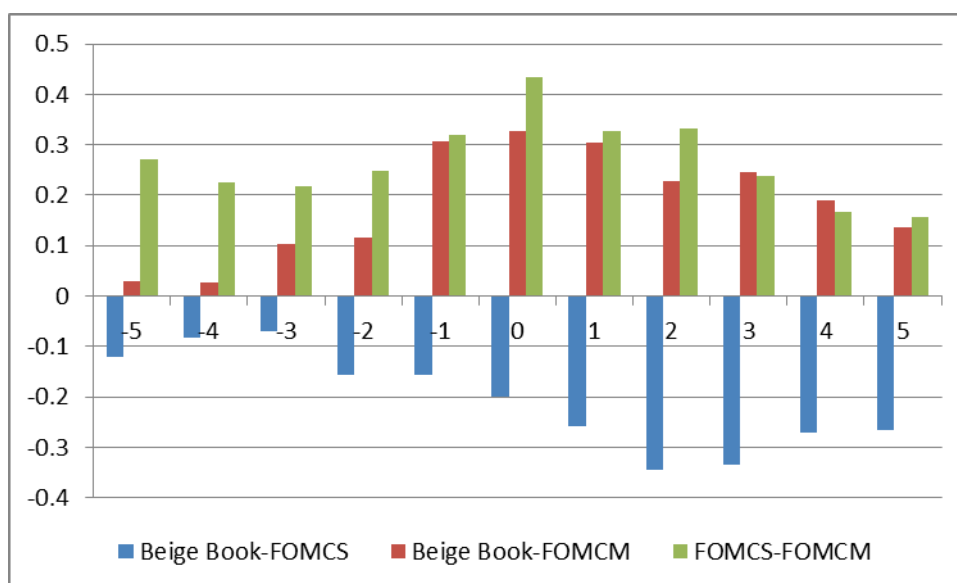
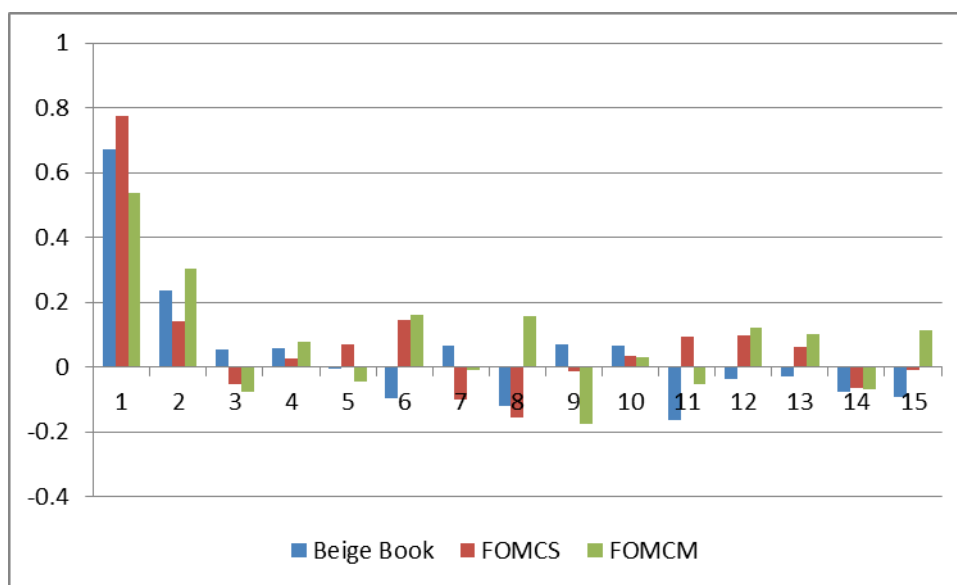
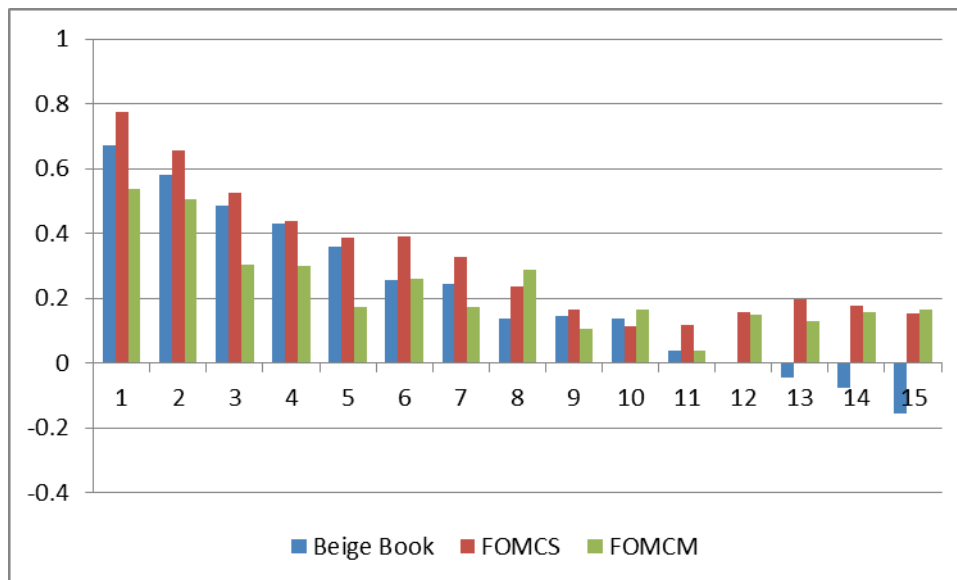


Figure 6. Top panel: Autocorrelations of the sentiment time series for the Beige Book, FOMC statements and minutes. The horizontal axis represents autocorrelation lag. Middle panel: partial autocorrelations of the sentiment time series for the Beige Book, FOMC statements and minutes. Bottom panel: cross-correlations for pairs of sentiment time series at different lags. The horizontal axis represents correlation lag. Correlations at lag 0 correspond to standard cross-correlations between the two time series (See Table 4).



Document	Frequency	Release time	Period from	to	# documents
FOMC Statement	8x per year		Feb 1, 2000	Mar 13, 2012	98
FOMC Minutes	8x per year	+ 3 weeks	Feb 1, 2000	Mar 13, 2012	98
Beige Book	8x per year	- 2 weeks	Jan 19, 2000	Feb 29, 2012	98

Table 1. The list of documents used in the analysis and the frequency of their publication. The release time shows when the given document is released relative to the FOMC statement.

Dependent variable: Length of Beige Book national summary

	<i>Coefficients</i>	<i>St. Dev.</i>	<i>t-statistic</i>	<i>p-value</i>	<i>R2</i>	<i>F-statistic</i>	<i>p-value</i>
b0	1888.76	100.01	18.89	0.00	0.30	22.11	0.00
b_trend	8.35	2.97	2.81	0.01			
b_Bern.	0.74	5.25	0.14	0.89			

	<i>Coefficients</i>	<i>St. Dev.</i>	<i>t-statistic</i>	<i>p-value</i>	<i>R2</i>	<i>F-statistic</i>	<i>p-value</i>
b0	1900.49	86.73	21.91	0.00	0.30	22.26	0.00
b_trend	8.00	2.01	3.98	0.00			
b_crisis	3.16	6.61	0.48	0.63			

Dependent variable: Length of FOMC Statements

	<i>Coefficients</i>	<i>St. Dev.</i>	<i>t-statistic</i>	<i>p-value</i>	<i>R2</i>	<i>F-statistic</i>	<i>p-value</i>
b0	229.28	18.87	12.15	0.00	0.67	99.68	0.00
b_trend	1.56	0.56	2.79	0.01			
b_Bern.	3.58	0.99	3.62	0.00			

	<i>Coefficients</i>	<i>St. Dev.</i>	<i>t-statistic</i>	<i>p-value</i>	<i>R2</i>	<i>F-statistic</i>	<i>p-value</i>
b0	222.98	15.60	14.30	0.00	0.70	115.00	0.00
b_trend	2.03	0.36	5.63	0.00			
b_crisis	5.86	1.19	4.93	0.00			

Dependent variable: Length of FOMC Minutes

	<i>Coefficients</i>	<i>St. Dev.</i>	<i>t-statistic</i>	<i>p-value</i>	<i>R2</i>	<i>F-statistic</i>	<i>p-value</i>
b0	5809.90	451.99	12.85	0.00	0.45	40.76	0.00
b_trend	-16.16	13.43	-1.20	0.23			
b_Bern.	119.45	23.74	5.03	0.00			

	<i>Coefficients</i>	<i>St. Dev.</i>	<i>t-statistic</i>	<i>p-value</i>	<i>R2</i>	<i>F-statistic</i>	<i>p-value</i>
b0	4924.89	423.21	11.64	0.00	0.36	28.38	0.00
b_trend	22.90	9.80	2.34	0.02			
b_crisis	93.72	32.23	2.91	0.00			

Table 2. The table presents three pairs of regressions of the trend with structural break for the length of the Beige Book, the FOMC minutes and the FOMC statements. The first regression features a structural break at the beginning of the Bernanke tenure, the second, at the beginning of the financial crisis during mid-2008. The two regression equations are respectively:

$$y = b_0 + (1 - D_{Bernanke}) * b_{trend} * t + D_{Bernanke} * b_{Bernanke} * t + e \text{ and}$$

$$y = b_0 + (1 - D_{crisis}) * b_{trend} * t + D_{crisis} * b_{crisis} * t + e,$$

where  $D_{Bernanke}$  (respectively  $D_{crisis}$ ) denotes the dummy variable equal to 0 prior to the Bernanke

tenure (prior to the crisis) and equal to 1 thereafter.

	T test				ANOVA	
	crisis	pre-crisis	t-value	p-value	F-statistic	p-value
Beige Book	5.48	5.51	-2.10	0.04	4.40	0.04
FOMC Statement	5.44	5.52	-5.31	0.00	28.22	0.00
FOMC Minutes	5.45	5.49	-7.35	0.00	54.06	0.00

Table 3. ANOVA analysis for the differences in the sentiment of the Beige Book, the FOMC minutes and the FOMC statements for the pre-crisis (before mid-2008) and financial crisis periods (mid-2008 and forwards). The tests appear to confirm that the average sentiment worsened during the financial crisis period.

Overall			
		Beige Book	FOMC statement
FOMC statement	correlation	-0.20	
	t-value	-1.99	
	p-value	0.05	
FOMC minutes	correlation	0.33	0.43
	t-value	3.40	4.71
	p-value	0	0

Greenspan tenure			
		Beige Book	FOMC statement
FOMC statement	correlation	-0.2004	
	t-value	-1.402	
	p-value	0.1674	
FOMC minutes	correlation	0.3859	0.2141
	t-value	2.868	1.503
	p-value	0.0062	0.1396

Bernanke tenure			
		Beige Book	FOMC statement
FOMC statement	correlation	-0.3978	
	t-value	-2.972	
	p-value	0.0047	
FOMC minutes	correlation	0.3181	0.1412
	t-value	2.3	0.9778
	p-value	0.0259	0.3332

Table 4. Correlations between sentiment ratings for the Beige Book and the FOMC minutes and statements.

swiss:finance:institute

c/o University of Geneva  
40 bd du Pont d'Arve  
1211 Geneva 4  
Switzerland

T +41 22 379 84 71  
F +41 22 379 82 77  
RPS@sfi.ch  
[www.SwissFinanceInstitute.ch](http://www.SwissFinanceInstitute.ch)