ECB Text Sentiment Analysis

# Executive summary

We performed text sentiment analysis on the ECB press conference speeches. By using four different sentiment dictionaries, we constructed four different sentiment score indicators. The scoring is done based on unigrams (individual words), with adjustment for negation words through bigrams (two adjacent words paired together). The sentiment indicator is then built up from the summation of the scores and its change measured across time.

# Methodology

## Preprocess: Tokenising words and attaching sentiment

We started by extracting ECB press conference transcripts since 2000, then converting transcripts into a list of individual words, in a process known as tokenising.

### List of top words and their associated sentiments

It is commonly used technique in text analytics to compare individual words against a dictionary of sentiments to determine the positivity of a given text. We use four dictionaries below to measure sentiment, as there are currently no publicly available dictionary specially for macroeconomic or central bank analysis.

|  |  |
| --- | --- |
| Dictionary | Description |
| AFINN | Scale of -5 (very negative) to +5 (very positive) |
| Loughran | For analysis of financial statements, with multiple types of sentiments |
| NRC | Crowd sourced emotion-based |
| Bing | Labels as positive or negative |

From an extract of the top words with associated sentiments, we see that the sentiments associated with each dictionary is quite varied.

*Table 1. Top words and their associated sentiments*

## # A tibble: 6 x 6  
## # Groups: word, n [6]  
## word n afinn bing loughran nrc  
## <chr> <int> <int> <chr> <chr> <chr>  
## 1 growth 4483 2 <NA> <NA> positive  
## 2 monetary 3760 NA <NA> <NA> anticipation|positive  
## 3 inflation 3396 NA <NA> <NA> fear|negative  
## 4 stability 2930 NA positive positive <NA>  
## 5 council 2738 NA <NA> <NA> anticipation|positive|trust  
## 6 policy 2676 NA <NA> <NA> trust

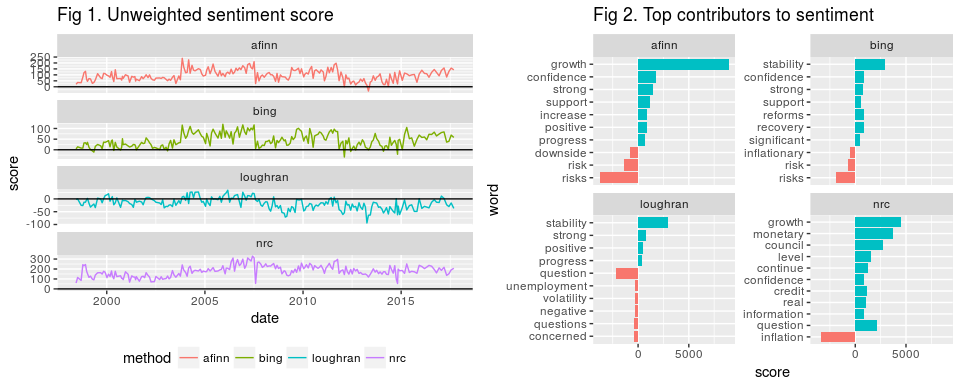
## Calculating a sentiment score

We then calculate a sentiment score by adding to the score when a positive word appears and subtracting when a negative word appears. In the case of the AFINN dictionary, we multiply the word count n by the sentiment score before summing them together. In Figure 2, we observe that each indicator is fairly different, suggesting that the dictionary has significant impact on your final results.

Using zero as a baseline, we observe that most of the dictionaries tend to give a positive overall sentiment to the minutes, but Loughran views the speeches negatively in most cases

### High scoring words from the dictionary

To account for the differences, we extracted the top contributors to scores both negative and positive across the dictionaries, and see the scores are attributable to fairly different words.



### Calculate weighted scores

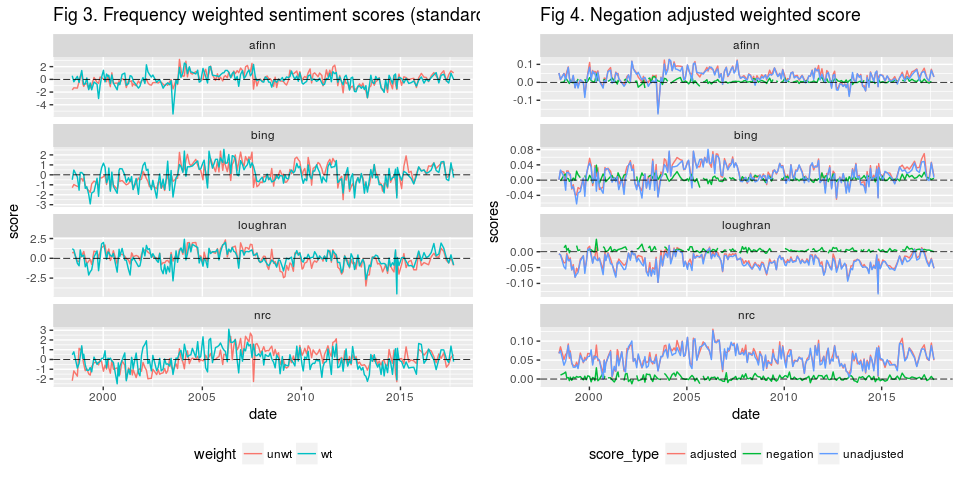
As words like growth or inflation may be used very often in any ECB speech, we may want to place more weight on words that occur less often. To adjust for this, we use "term frequency - inverse document frequency" (tf-idf) as a weight. tf-idf measures how unique a particular word is in the other documents. If it is relatively unique, it is given a higher weight, while if it appears frequently in every single document, it would be given a smaller weight. Using the tf-idf as weights, we calculated a weighted score in Figure 3.

## Adjusting for bigrams

### Removing negation words

Unigram analysis may not be sufficient, as there are words which are used with opposite intentions, such as "negative growth" or "no money". Hence we extend the analysis to bigrams, which are essentially words paired together with their adjacent words.

To adjust the negated words, we have to first remove its original impact, then add the new impact, which is equivalent to adding twice its new established impact. From Figure 4, we observe that negation has had very small impact on the overall sentiment score, which suggests that in central bank speeches, negation terms are not significantly used.

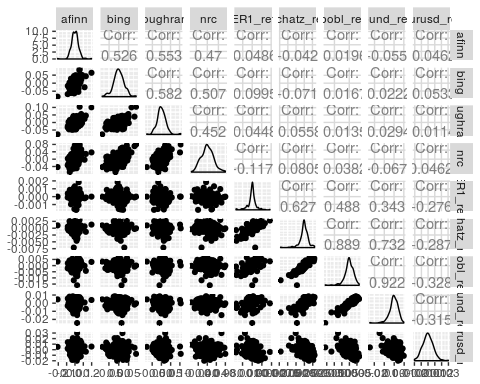


## Predictive ability of sentiment

To test our sentiment scores, we compared the indicators against 1 day performance of Euribor and bond futures on the day of the announcement.

Performance of the sentiment indicators as market indicators is poor. If we look at the correlation between each of the indicators and the performance of the assets, the correlation is consistently less than 10%

*Fig 5. Correlation between weighted indicators and 1 day return on assets*



# Limitations of study

There are several limitations in this exploratory study of text sentiment analysis.

1. The first, and biggest issue is the dictionaries used. As macroeconomics and central bank speech can have very unique terminologies, or against more generalized financial terms. As there are no readily available dictionaries for macroeconomic analysis at the moment, having one could improve performance significantly.
2. Market movements are a result of market expectations vs central bank stance, hence instead of levels of positivity in the speech, it may be more useful to compare relative levels of positivity between news and market commentary leading up to the meeting, and the eventual tone of the meeting itself. However, much more data will need to be collected to discern market sentiment.

# Evaluation

There is still a long way to go for us to effectively use text analytics to generate signals. However, if the limitations can be improved upon, there is potential for the indicator to help glean objective information from news sources or central bank speeches.

Text sentiment analysis can also be applied in different ways such as picking out words that might be associated with hawkish or dovish moves by comparing words used to meetings where measures were announced.