Imperial College London
BUSINESS SCHOOL

Macro Finance

Coursework 2

20th February 2024

Group 1

STUDENT NAME	CID NUMBER
George Kapadoukakis	02477138
George Valsamis	02371460
Kiyana Amiri	02444882
Long Him Lui	01853906
Sebastiano Raffaelli	02527205

Table of Contents

Table of Figures

Introduction

Using Python, we analyzed the 81 text files used by the Federal Open Market Committee (FOMC). We compared two libraries, namely the LM Dictionary and the Harvard Dictionary. Initially, the project required us to clean our data by removing capital words, making sure words like "Cross" and "Crossing" were grouped, eliminating filler words like "and" from our analysis, and tokenizing the files by splitting text strings into individual tokens to separate sentences and turn them into unigrams. We then started our investigation by counting the frequency of each word. This gave us a rough idea of the most used words. We then calculated the tone for each document, with 0 being a negative tone and 1 being a positive tone. For visual purposes, we plotted the 81 tones on a time series against the SP500 for comparison to see if there is a relationship between negative tones and downward movements in the SP500 and vice versa. We also graphed the polarity of the text over time, the polarity being defined in the "pysentiment2" library (Python Sentiment Analysis) as:

$$Polarity = \frac{Positive\ Frequency + Negative\ Frequency}{Total\ Words}$$

Additionally, we wanted to compare the LM and Harvard Library by counting the number of negative and positive words they had observed throughout time. This would allow us to spot differences between the two libraries. On top of this, we plotted these differences with one graph representing positive frequencies and the other representing negative frequencies. To compare the dictionaries even further, we separated the words into negative, positive, and neutral, and we created a table showing the most used positive and negative words in both libraries. Finally, we performed two regressions against the SP500 as a dependent variable, using the LM Dictionary and Harvard Dictionary tones as independent variables.

Question 1.1

In the first section, the analysis focuses on quantifying the frequency of terms in the provided FOMC documents, aiming to find the most influential unigrams that characterize the monetary policy communications of the FED during 2008-2018.

The methodology starts by processing the text by removing punctuation and digits, followed by tokenization to split the document into individual unigrams. We also filtered out common stopwords such as "the" and "is", which provide little value in assessing the tone of the documents.

The results are displayed below with in Figure 1:

Word	Frequency
committee	927
inflation	670
economic	497
federal	411
rate	390
market	357
policy	339
securities	308
conditions	302
percent	272

Figure 1: FOMC Most Influential Unigrams

When examining the most frequent words, they align with the purpose of the FOMC documents and monetary policy, which are to maintain price stability, stabilizing inflation, as well as maximizing employment. The words "committee", "inflation" and "policy" reflect key areas of focus for the FED, which highlights the role of the institution regarding inflation, economic health, and policy direction. An important point to note would be the prevalence of the words "federal" and "rate", which showcases the importance of the federal fund rate in the FOMC and FED's actions.

Question 1.2

The second section asks us to quantify the tone of the FOMC documents provided from the period 2008-2018, where we utilize sentiment analysis tools within Python, such as NLTK's Sentiment Intensity Analyzer and PySentiment 2 library to quantitatively evaluate the tone.

We present our results below in Figures 2 and 3:

Date	Tone
2008-01-22	-0.034483
2008-01-30	-0.016807
2008-03-11	0.008547
2008-03-18	-0.036496
2008-04-30	-0.021277
…	…
2017-07-26	0.015625
2017-09-20	0.000000
2017-11-01	0.000000
2017-12-13	0.011236
2018-01-31	0.032051

Figure 2: Tone Measurements

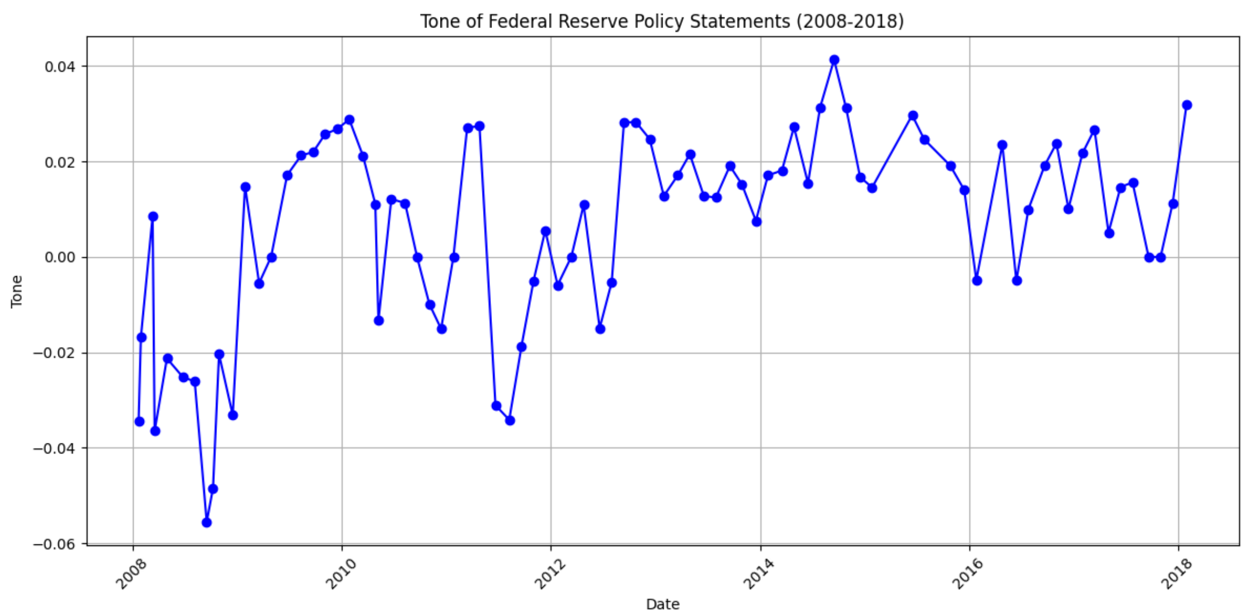

Figure 3: Tone of FOMC Statements (2008-2018)

Our results reveal a dynamic interaction between the FOMC documents, signifying central bank communications, and market developments. The period covers the aftermath of the 2008 financial crisis and its subsequent recover phases, which offers room for analysis into the FED's stance on monetary policy through turbulent economic conditions.

In the immediate aftermath of the 2008 financial crisis, we can see that the FOMC documents had a more cautious tone, which showcases the severity and uncertainty of the economic conditions in the early recovery phases. History shows that this period included many unprecedented policy interventions such as quantitative easing and zero-interest-rate policies with the aim of stabilizing financial markets and promoting economic recovery. The more negative tone analyzed from the documents signals the FED's commitment to use the tools at their disposal to support the economy and provide reassurance to the public.

From the period of 2008 onwards, we can observe the FOMC documents shifting to a more optimistic tone, which was particularly evident in 2015, when the FED raised the interest rates for the first time since the crisis. This illustrates how the tone of the FOMC document can be used to manage expectations and minimize market disruptions.

During the period, we can also see periods of volatility. Some notable ones being the European debt crisis in 2011 and the Taper Tantrum in 2013. During these periods, we can see the FED's tone switching between expressing concern over global economic risks and reaffirming confidence in economic recovery. Therefore, this showcases the times FOMC document tone has helped dampen market overreactions and highlights the importance of the FED's role as a stabilization mechanism.

Question 1.3

This question shifts the focus to comparing the LM and Harvard dictionaries when measuring the tone of the FOMC documents, specifically FOMC minutes. We quantify this by separating the words into positive and negative, which is classified by each dictionary, and quantifying each classification.

Our results are displayed in Figure 4,5 and 6 below:

Document Date	LM Positive	LM Negative	Harvard Positive	Harvard Negative
2008-01-22	0	4	19	8
2008-01-30	0	2	19	7
2008-03-11	1	0	25	3
2008-03-18	1	6	26	13
2008-04-30	2	5	31	11
…	…	…	…	…
2017-07-26	6	3	47	18
2017-09-20	7	7	48	24
2017-11-01	6	6	44	24
2017-12-13	6	4	40	22
2018-01-31	6	1	35	16

Figure 4: Positive and Negative Unigram Count for LM and Harvard Dictionaries

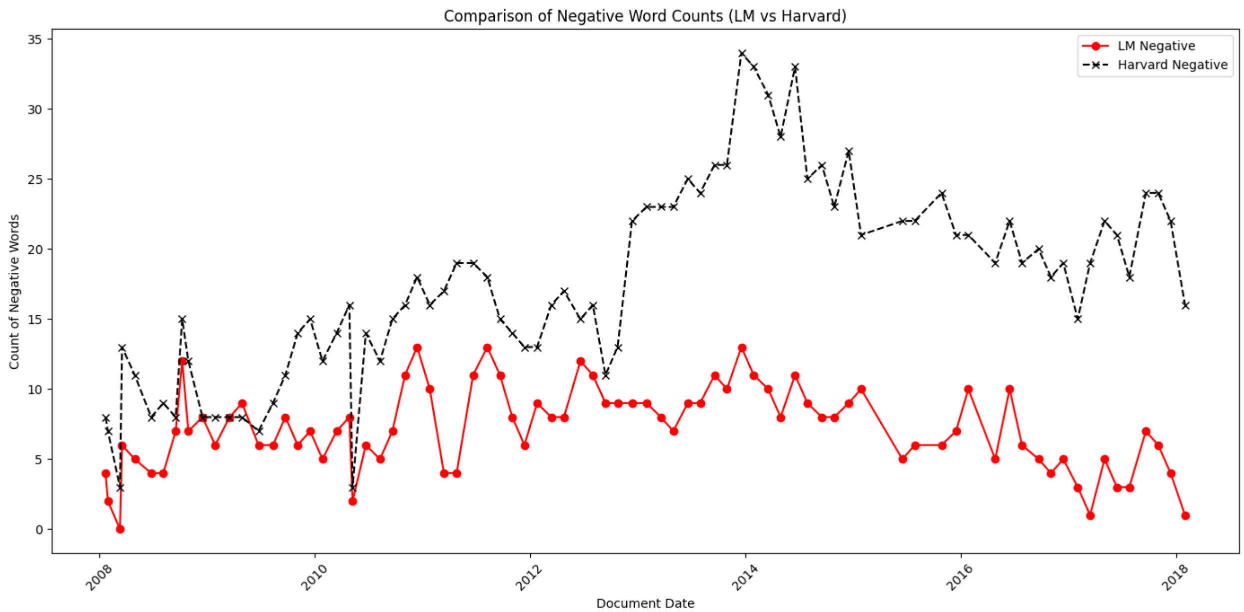

Figure 5: Comparison of Negative Words Harvard vs. LM

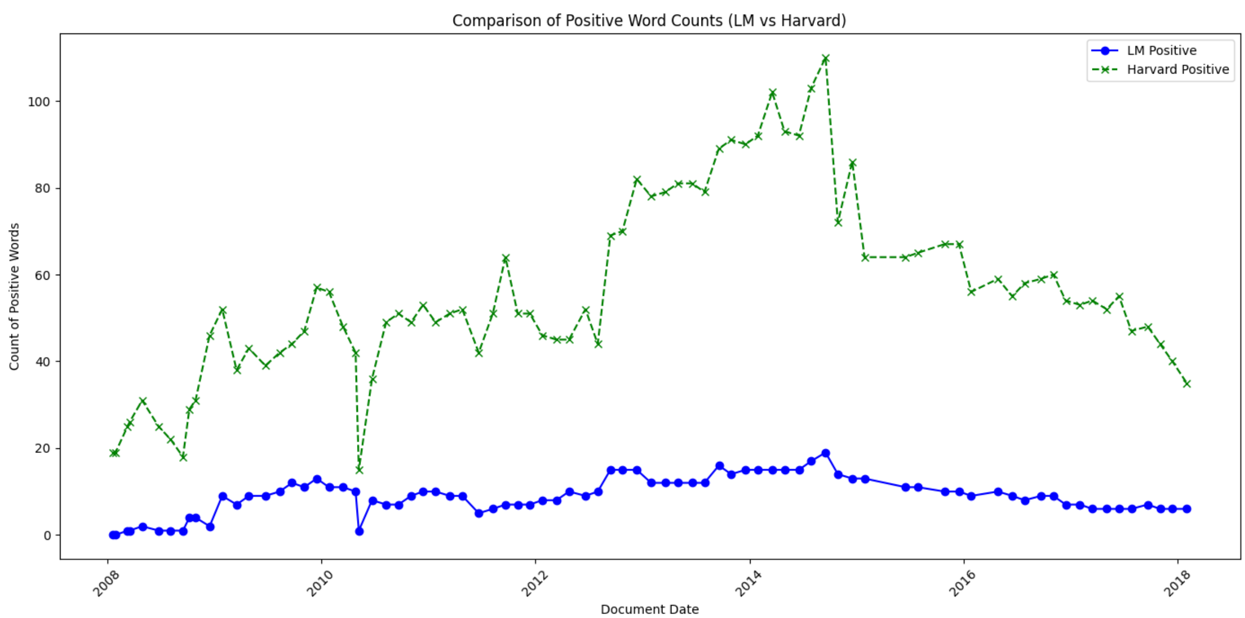

Figure 6: Comparison of Positive Words Harvard vs. LM

Top 10 Positive Words:

Word	Frequency
stability	162
improvement	75
progress	74
stable	61
exceptionally	57
stronger	39
improved	37
strengthen	33
gains	22
strengthens	21

Figure 7: Top 10 Positive Words LM

Top 10 Negative Words:

Word	Frequency
unemployment	97
declines	36
declined	32
decline	26
weak	25
slowed	23
depressed	22
downward	21
slow	19
strains	17

Figure 8: Top 10 Negative Words LM

Top Positive Words:

	Word	Count
0	securities	308
1	growth	135
2	support	97
3	treasury	95
4	help	85
5	energy	76
6	improvement	75
7	progress	74
8	asset	68
9	stable	61

Figure 9: Top 10 Positive Words Harvard

Top Negative Words:

	Word	Count
0	low	115
1	unemployment	97
2	risks	88
3	debt	61
4	pressures	59
5	lower	33
6	downside	30
7	weak	25
8	depressed	22
9	pressure	20

Figure 10: Top 10 Negative Words Harvard

It is important to first understand some context behind the two dictionaries. The LM dictionary is specifically designed for finance and economic texts and has a tailored approach to analyzing sentiment towards those fields. On the other hand, the Harvard dictionary encompasses a broader approach to sentiment analysis, covering a wide range of positive and negative sentiments. However, this may not capture specialized language as precisely as the LM dictionary.

Our findings suggest that the LM dictionary may provide a more detailed understanding of the FED's policy stance. This may be due to its more focused emphasis on financial and economic applications, which yields a more nuanced reflection of sentiment. We can see that the frequency of both positive and negative terms algin closely with the key economic and policy-related topics mentioned in the FOMC minutes. This alignment showcases the LM dictionary's specialized focus and ability to capture the sentiment of monetary policy documents, being able to differentiate the specific connotations with an economic context.

To contrast this, the Harvard dictionary does not offer the same level of detail as the LM dictionary with respect to the financial domain. While the Harvard dictionary was able to identify general sentiment trends within the FOMC documents, its broader categorization of positive and negative words may not capture the language present in monetary policy articles. This observation is supported by the comparative counts of positive and negative words from the Harvard dictionary compared to the LM dictionary. Despite being able to indicate similar overarching trends, the Harvard dictionary lacks the precision offered by the LM dictionary's economic applicability.

By conducting further research on the two libraries, we gain even more insight into our findings. The LM dictionary takes economic context into account, which assigns particular words different sentiment values in the context of monetary policy rather than what is inferred from general language use. For instance, terms like "inflation" and "growth" may carry different implications in the context of monetary policy compared to their sentiment values, such as if there is "negative inflation" or "negative growth".

Therefore, our conclusion from using the two libraries to analyze the tone of the FOMC documents showcase the LM library's superiority to the specific economic-related application.

Question 1.4

The following section evaluates the impact of the FOMC documents through the lens of the S&P 500 index performance. We perform regression techniques to correlate sentiment scores for both the LM and Harvard dictionaries against S&P 500 stock market returns.

The results from the are displayed below in Figures 11 and 12:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                Returns   R-squared:                       0.094
Model:                            OLS   Adj. R-squared:                  0.082
Method:                 Least Squares   F-statistic:                     8.149
Date:                Mon, 19 Feb 2024   Prob (F-statistic):            0.00550
Time:                        23:51:56   Log-Likelihood:                 139.25
No. Observations:                  81   AIC:                            -274.5
Df Residuals:                      79   BIC:                            -269.7
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0020      0.005      0.394      0.694      -0.008       0.012
Tone           0.6777      0.237      2.855      0.006       0.205       1.150
==============================================================================
Omnibus:                        6.398   Durbin-Watson:                   1.972
Prob(Omnibus):                  0.041   Jarque-Bera (JB):                7.730
Skew:                          -0.339   Prob(JB):                       0.0210
Kurtosis:                       4.353   Cond. No.                         48.7
==============================================================================
```

Figure 9: LM Regression Statistical Results

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                Returns   R-squared:                       0.023
Model:                            OLS   Adj. R-squared:                  0.011
Method:                 Least Squares   F-statistic:                     1.876
Date:                Mon, 19 Feb 2024   Prob (F-statistic):              0.175
Time:                        23:52:56   Log-Likelihood:                 136.23
No. Observations:                  81   AIC:                            -268.5
Df Residuals:                      79   BIC:                            -263.7
Df Model:                           1
Covariance Type:            nonrobust
================================================================================
====
                       coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
----
const               -0.0296      0.027     -1.107      0.272      -0.083
0.024
Harvard_Polarity     0.0694      0.051      1.370      0.175      -0.031
0.170
==============================================================================
Omnibus:                       14.446   Durbin-Watson:                   1.677
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               18.368
Skew:                          -0.821   Prob(JB):                     0.000103
Kurtosis:                       4.657   Cond. No.                         12.7
==============================================================================
```

Figure 10: Harvard Regression Statistical Results

Firstly, we can examine the coefficients derived from the regression model, which quantify the relationship between the sentiment scores and S&P 500 returns. The positive coefficient from the LM dictionary's sentiment score suggests a direct relationship where a more optimistic tone in the FOMC documents leads to an uptick in S&P 500 returns, which favors the LM dictionary over the Harvard dictionary.

Secondly, the p-values assess the probability that the observed relationship has occurred by chance. The LM dictionary has a low p-value, which affirms the statistical significance of its sentiment scores on S&P 500 returns. However, the higher p-value for the Harvard dictionary suggests its sentiment analysis does not have as much statistically significant predictive power on S&P 500 market returns.

Furthermore, the F-statistic can be employed to evaluate the overall fit of the regression model by testing the null hypothesis that all regression coefficients are equal to zero. The higher F-statistic from the LM dictionary sentiment scores compared to the Harvard dictionary indicates better overall fit and explains a greater proportion of variance in the returns.

The T-statistic determines the statistical significance of each independent variable within the model. The higher T-statistic for the LM dictionary's model compared to the Harvard dictionary's model indicates that the sentiment scores for LM are better significant predictors of S&P 500 returns.

However, arguably the most important measure is the R-Squared measure, which evaluates the proportion of the variance in the S&P 500 returns that is predictable from the sentiment scores. The higher R-Squared measure of the LM dictionary's model compared to the Harvard model suggests that the LM model has stronger explanatory power.

To conclude, the above statistical measures provide a quantitative analysis of the predictive power of sentiment analysis of the two libraries on S&P 500 market returns and showcases the LM dictionary's superior capability given financial contexts.