Decoding Fed: A Machine Learning Approach on Deciphering FOMC Meetings

Haoze Song

Apr 30, 2018

Introduction

Nowadays there are countless factors that are closely related to the stock market. As financial analysts on Wall Street, we are interested in finding out their correlations with the change in market to predict market trends. In the report, we aim to how the monetary policies implemented by the Federal Reserve (Fed) can reflect the current stock price. Specifically, we apply Natural Language Processing (NLP) techniques to the meeting minutes of the Federal Open market Committee (FOMC) in our analysis.

The FOMC, which consists of twelve members, "holds eight regularly scheduled meetings per year. At these meetings, the Committee reviews economic and financial conditions, determines the appropriate stance of monetary policy, and assesses the risks to its long-run goals of price stability and sustainable economic growth."[1] These meeting minutes are released by FOMC about three weeks after the meetings, which reveal to the public FOMC's view of the economy and its monetary policy decisions.

During an FOMC meeting, the Fed discusses a wide range of topics such as interest rates, employment, and prices of goods and services. These topics impact the financial market in different ways, and consequently, one of our goals in the project is to identify the topics from the FOMC meetings and to examine their correlations with the market. To accomplish our goal, we apply the Latent Dirichlet Allocation (LDA) model to the FOMC meeting minutes to automatically identify topics. After the topic retrieval, we calculate the percentage change of topic weights in each minute overtime and analyze how the change relates to the macroeconomic conditions. We also analyze the sentiments of each topic and see how they vary overtime and build models that relate the sentiment change with the market performance.

The rest of the report will describe in detail the techniques and models we use in the project. We also include some visual representations to showcase our findings.

Data

We obtain and compile the FOMC minutes data ranging from 1990 to present from the Fed official website. To facilitate our model implementation, we organize the data by paragraph. Since the minutes is presented in different formats year by year, we write various versions of Python Beautiful Soup algorithm to extract and rearrange the data. By deleting information that is of less interests (FOMC voting results, members introduction, operation details, etc.),

[1]Source: https://www.federalreserve.gov/monetarypolicy/fomc.htm

we select and preserve 9,125 unique paragraphs across the target timeframe that we consider most relevant to macroeconomic forecasts and viewpoints for our analysis.

Topics Identification

We apply the Latent Dirichlet Allocation (LDA) algorithm, which is an unsupervised machine learning model developed by Blei, Ng, and Jordan (2003) to classify the text into various distinct themes. Under the LDA setting, if we assume each of paragraph *d* contains a mixture of *N* topics, and the proportion of topic *n* in this paragraph is $\theta_{n,d}$. Let the vector $\theta_d = [\theta_{1,d,}\theta_{2,d}, \dots, \theta_{N,d}]$ represents the true topic mixture distribution of a given paragraph, this vector should follow an N-order Dirichlet distribution specified by the latent parameter of vector μ:

$$\theta_d \sim \ Dirichlet_N(\mu)$$

For each topic *i*, let β_i be the vector of probabilities of observing word *j* in topic *i* and let the vector $\beta = [\beta_1, \beta_2, \dots, \beta_N]$ represents the collection of all the topics. Each β_n should follow a Dirichlet distribution governed by the latent parameter ϕ:

$$\beta_n \sim \ Dirichlet(\phi)$$

Let the assignment of topics to word *i* in paragraph *d* be $Z_{d,i}$. We assume it follows a multi-nomial distribution specified by the parameter θ_d that we have just illustrated:

$$Z_{d,i} \sim Multinomial(\theta_d)$$

Correspondingly, for each word *i* in paragraph *d*, we assume the assignment according to the chosen topic, $W_{d,i}$, follows a multi-nomial distribution governed by the resulting word-topic assignment parameter $\beta_{Z_{d,i}}$:

$$W_{d,i} \sim Multinomial(\beta_{Z_{d,i}})$$

The four parameters developed above describe the latent data generating process, which is given by the following joint distribution of all the latent variables:

$$p(\beta_{1:N}, \theta_{1:D}, Z_{1:D}, W_{1:D}) \\ = \prod_{n=1}^{N} p(\beta_i) \prod_{d=1}^{D} p(\theta_d) \, [\prod_{i=1}^{I_d} p\big(Z_{d,i}|\theta_d\big) \, p(W_{d,i}|\beta_{1:N}, Z_{d,i})]$$

Using the Baye's rule, we can then calculate the distribution of the intended topic-paragraph assignment:

$$p(\beta_{1:N}, \theta_{1:D}, Z_{1:D}|W_{1:D}) = \frac{p(\beta_{1:N}, \theta_{1:D}, Z_{1:D}, W_{1:D})}{p(W_{1:D})}$$

We run the LDA algorithm using R on the data set we obtain based on the previous steps. To select the optimal number of topics, we run the model from 5 to 10 topics. Since the FOMC meeting is held 8 times a year and every time there is a specific focus, we decide to choose 8 as our tipic numbers. The model perplexity, at the same time, also negatively correlated with the number of topics, according to the figure below. In order to balance between model perplexity and potential risk of over-fitting, 8 is an appropriate number to apply.

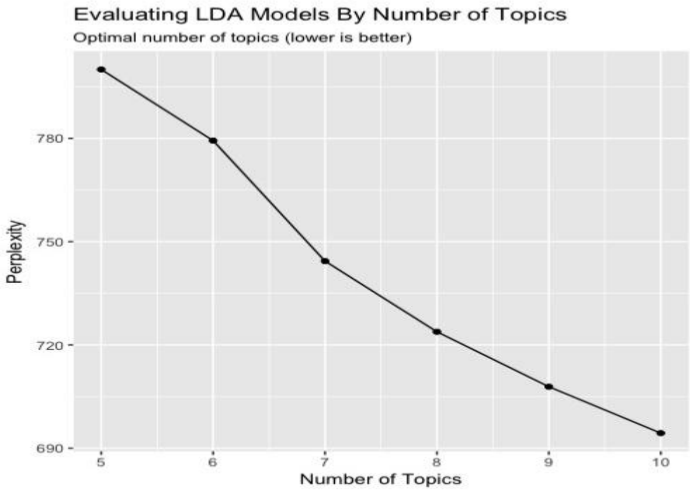

Figure 1. Model Perplexity & Number of Topics

The distribution of each keyword according to different topic is presented in the following figure. According to the keywords identified, we make detailed interpretation of each topic, and summarize our topic description in the table below.

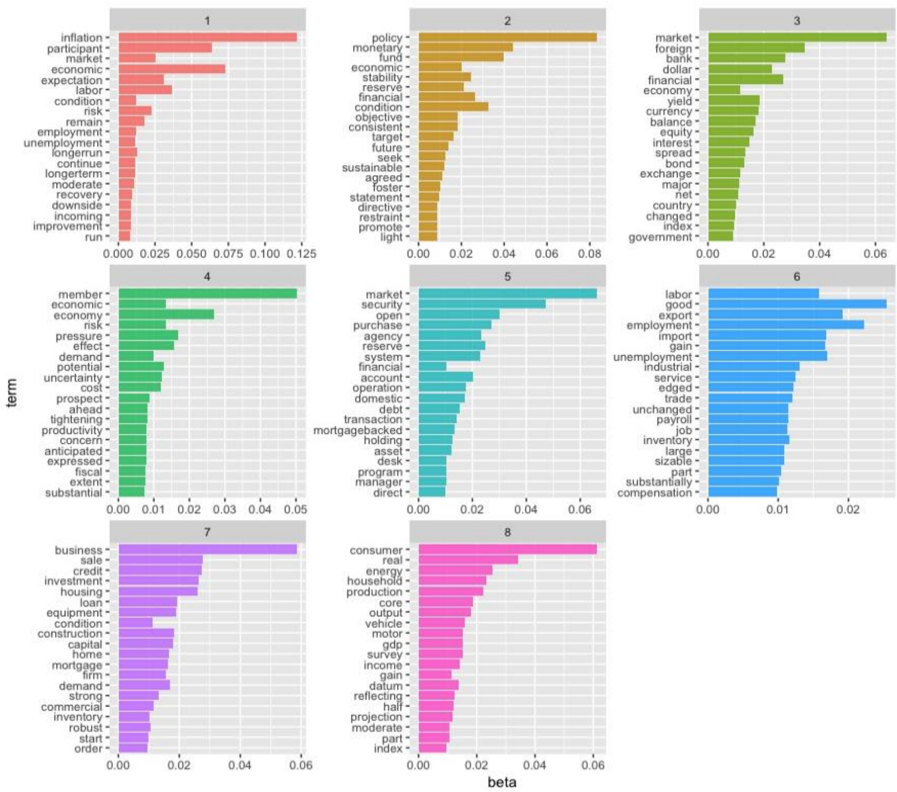

Figure 2: Distribution of Top LDA Keywords

	Topic Interpretation
1	Inflation
2	Policy
3	Financial Markets
4	Economic Growth
5	Open Market Operation
6	Employment
7	Investment
8	Production & Consumption

We find that the proportion of each topic varies significantly over time. Such change of Fed focus is caused by various reasons: market conditions instability (2008-2009), temporary change of Fed reporting subjects, or significant geopolitical events. The change of such portion is illustrated in the following figure (recession period has been shaded). We also generate a Moving Average figure to make it easier to identify the trends.

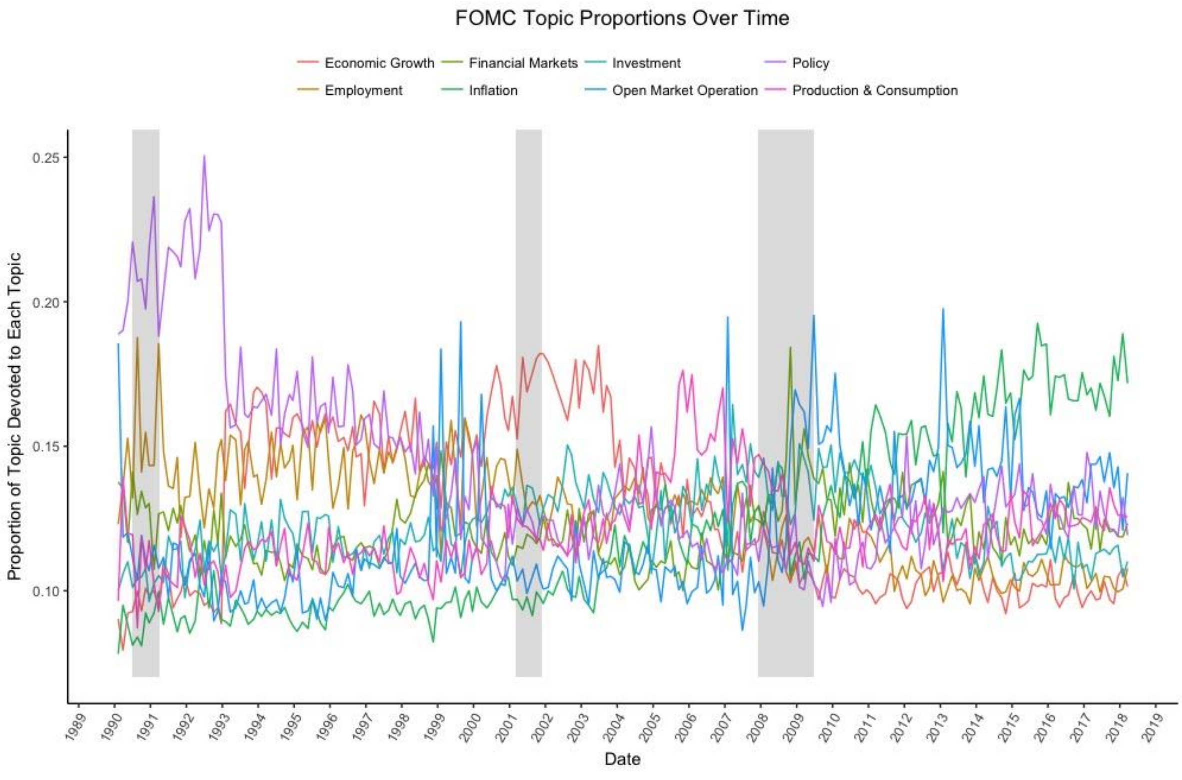

Figure 3. FOMC Minutes Topic Proportion Over Time (Recession shaded)

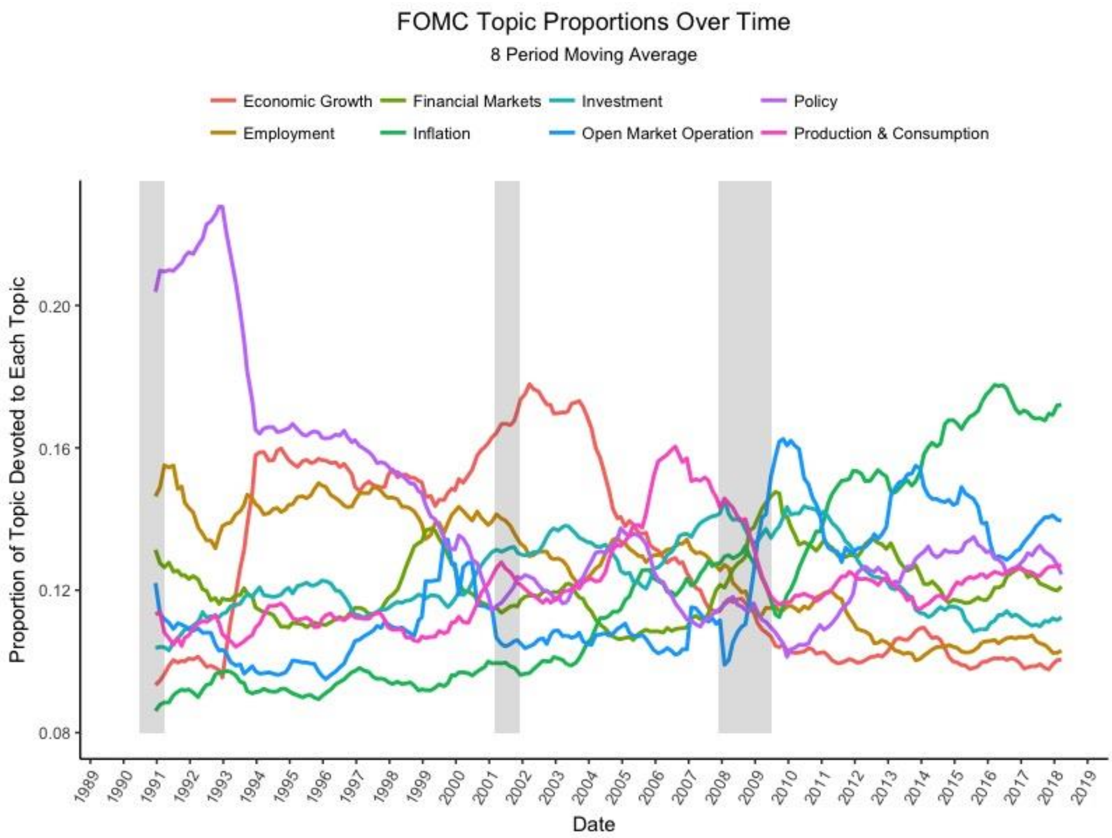

Figure 4: FOMC Minutes Topic Proportion Over Time (Recession shaded)

Various interesting patterns can be identified in this plot. From 1990 to present, we notice that Fed has significantly increased their focus on inflation, employment and open market operation, which corresponds to the central bank's efforts in devoting to the "dual mandates": maximizing employment and stabilizing price level. During the 2008-2009 Global Financial Crisis, we found that Fed boosted their rhetoric around open market operation and financial markets, which corresponds to their macroscopic adjustment and control to stabilize the market-wide turbulence. The emphasis on economic growth has increased since 1992, peaked around 2002, and gradually decreased over the years. The significant variation of such topic proportion corresponds to the presidency of Bill Clinton, who presided over the longest period of peacetime economic expansion in American history. The Congressional Budget Office also reported budget surplus during that last three years of his presidency. After 2002, the focus on US gradually changed to global counter-terrorism after the "9-11" attacks. Such subtle and vital change of US macro policies is well reflected from the FOMC topic proportion variation.

Sentiment Analysis

We now try to analyze the sentiments of each topic in the meeting minutes. Specifically, we first use the comprehensive tonal lexicon developed by Loughran and McDonald (2011) to determine the tone of each word in those minutes. For each paragraph in a minute, we count the number of positive-tone keywords and negative-tone ones according to the Loughran Dictionary and score the paragraph as the positive counts minus the negative counts. The score is used to decide the tone of the paragraph, with a higher score indicating a more positive/easing position and a lower score indicating a negative/tightening position. Within each paragraph, we assign the score of a particular topic to the paragraph score multiplied by the weight of the topic given by LDA in the previous section. We then aggregate the topic scores to the document level as the weighted sum of the paragraph scores, where each weight is the inverse of the number of words in the corresponding paragraph.

$$Score_{i,t} = \sum_{d=1}^{D_t} Score_{d,i,t}\left(\frac{1}{T_d^t}\right),$$

where $Score_{i,t}$ is the tone score for topic i in Minute t, $Score_{d,i,t}$ is the tone score for topic i in paragraph d of Minute t, T_d^t is the total word counts in paragraph d in Minute t, and D_t is the total number of paragraphs in Minute t.

Figure 5 plots the document-level score for the tone of the eight topics over time. We use the standardized scores for each topic score time series so that it is easier to make comparisons. The gray areas correspond to the recession period in our analysis. The figure suggests that during recessions, the general tone of each topic becomes more negative, which caters to our expectations. It also shows that the tone of each topic tends to move together. We spot that the sentiment score for the topic Investment deviates from the others from 1994 to 1998. That might happen because the economy was booming in the mid-90s, so the Fed had a positive outlook on business investment for the period.

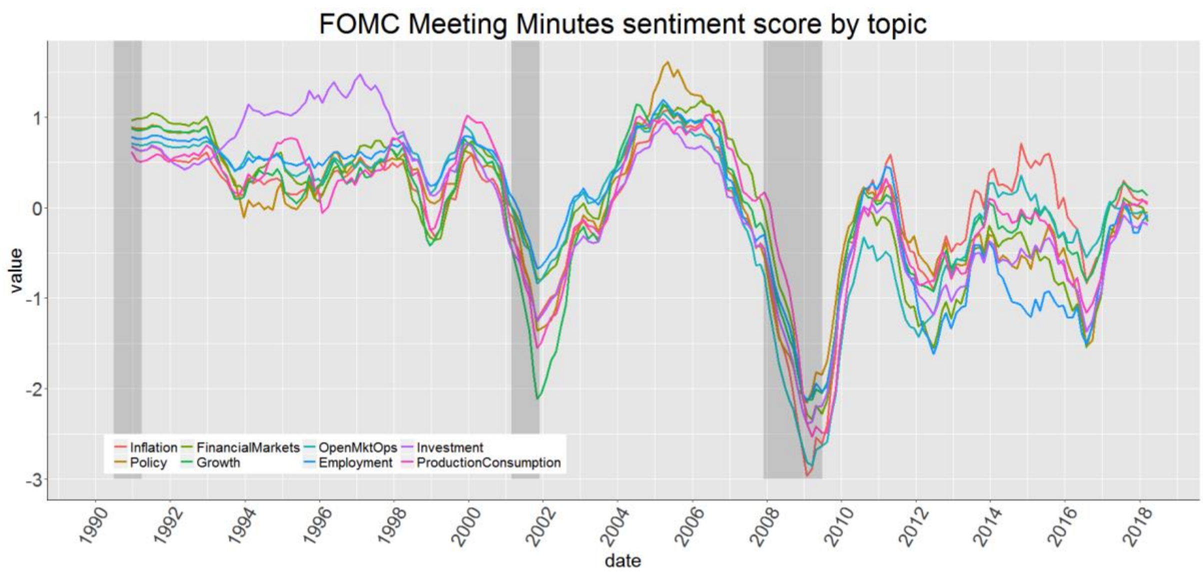

Figure 5: FOMC Minutes Topic Tone Scores Over Time (Recession shaded)

Regression Analysis

We seek to find some relationship between the topic proportion and the macroeconomic conditions for each of the above eight topics. We use interest rate, unemployment rate, and whether the date is in a recession as factors for the macroeconomic conditions, whose data are obtained from different sources. Specifically, the interest rate is extracted from the US Treasury website, the unemployment rate is found from the Bureau of Labor Statistics, and the recession classification is decided depending on whether the meeting date is within a NBER-designated recession period. We fit a linear regression model to all of our meeting minutes data to examine the relationship:

$$\hat{p}_{i,t} = c + a1 * InterestRate_t + \ a2 * UnemploymentRate_t + a3 * Recession_t + \ e_t \ ,$$

where $\hat{p}_{i,t}$is the proportion for topic i in minutes t from the LDA process. The following table presents the coefficients estimate. Note that the OMO in the table stands for Open Market Operations and the Pro & Con means Production and Consumption.

Topics								
	Inflation	Policy	Market	Growth	OMO	Employment	Investment	Pro & Con
Interest Rate	-0.0133 ***	0.0125 ***	-0.0012 **	0.0044 ***	-0.0060 ***	7.6e-03 ***	-0.0011 *	-0.0030 ***
Unemploy Rate	-0.0009	0.0039 ***	0.0023 ***	-0.0060 ***	0.0032 ***	-9.7e-04 *	0.0014 *	-0.0029 ***
Recession	-0.0069 *	-0.0060	0.0115 ***	-0.0080	0.0057	-7.6e-05	0.0034	0.0004
Adj. R-squared	0.7087	0.505	0.2477	0.2338	0.3259	0.678	0.06	0.18

From the table, we indeed find some relationship between the topic proportions and the macroeconomic conditions. For instance, the proportion of inflation-related topic is negatively related to interest rate, unemployment rate, and the recession. However, several of the Adjusted R-squared of our eight linear regression models are very small, indicating the inadequacies of the models. Finding the relationship between the topic proportion and the macroeconomic conditions is not the main purpose of our report, so here we are just building the regression model using the most basic factors. A more complicated selection of factors to represent the macroeconomic conditions will certainly enhance the result.

We also seek to examine the relationship between the topic tone score and the macroeconomic conditions. The set up of our model is very similar to the one above, where we fit a linear regression model:

$$Score_{i,t} = c + a1 * InterestRate_t + \ a2 * UnemploymentRate_t + a3 * Recession_t + \ e_t.$$

The table on the next page presents the coefficients estimate:

Topics								
	Inflation	Policy	Market	Growth	OMO	Employment	Investment	Pro & Con
Interest Rate	0.1685 ***	0.2465 ***	0.3016 ***	0.2213 ***	-0.2390 ***	0.3125 ***	0.3322 ***	0.2393 ***
Unemploy Rate	0.0011	0.0165	-0.0531	-0.0048	-0.1100 ***	-0.0440	-0.0085	-0.0602
Recession	-1.5681 ***	-1.3067 ***	-1.1450 ***	-1.4936 ***	-1.5623 ***	-1.1450 ***	-1.4560 ***	-1.5043 ***
Adj. R-squared	0.3159	0.3455	0.4478	0.3624	0.4771	0.4657	0.562	0.4218

From the table, that influence of the unemployment rate in contributing to the topic tone score is not significant compared to the other two factors, whereas that of the other two factors is very significant. This might happen because the information given by the unemployment rate is already included in that of the Interest Rate and Recession factors.

Conclusion

In this paper, we explore, implement and present a machine learning approach in understanding Fed FOMC meeting minutes. More specifically, we apply the LDA algorithm to determine the 8 topics discussed in the minutes, namely Inflation, Policy, Financial Markets, Economic Growth, Open Market Operation, Employment, Investment, Production & Consumption. We study the variation of proportion of these topics over time and discuss the correspondence between such changes and deep-level US macroeconomic changes. In addition, we conduct sentiment analysis of the tone of Fed speak, and conclude that the tone tends to be negatibe during recession and tends to be positive when economy is booming. Such finding corresponds to our intuition. Finally, we implement 2 linear regression models: one is between our identified topic proportions and macroeocnomic variables, the other is between minuets sentiment scores and macroeconomic variables. Based on the fitting results, we prostulate that certain topics are comparatively more informative than others (Inflation, Employment). To further validate our findings, researching on the relationship between the newly-released minutes contents and instantaneous price changes, or volatility level of financial assets might be a fruitful direction. Advanced algorithmic trading strategies can be developed to exploit such subtle connection. Due to timing constraints for this project, we leave this for future research.

References

Blei, Ng & Jordan (2003). *Latent Dirichlet Allocation*. Journal of Machine Learning Research. 3(Jan):993-1022, 2003.

Fed (2018). *About the FOMC*. Federal Open Market Committee. Retrived 13th Apr 2018 from https://www.federalreserve.gov/monetarypolicy/fomc.htm

Loughran & *McDonal*d (2011). *Documentation for the Loughran McDonald Master Dictionary*. Retrieved 13th Apr 2018 from https://www3.nd.edu/~mcdonald/Word_Lists_files/Documentation/Documentation_LoughranMcDonald_MasterDictionary.pdf