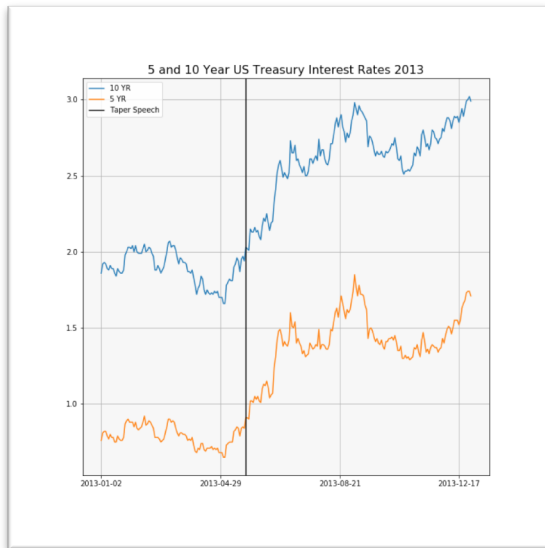


Background:

Federal Reserve is responsible for setting short term interest rates and the money supply in an attempt to keep inflation and unemployment within an acceptable range. Federal Reserve Board Member speeches and press releases by the committee that sets interest rate policy may provide insights into future changes in interest rates.



The Goal:

Forecast future changes in interest rates based on Euclidean distance between current Federal Reserve text and recent text.

The Process:

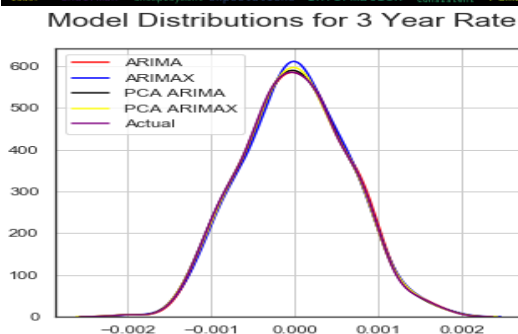
Two different data sets were tested.

- All Federal speeches from 2006 to present
- FOMC minutes from 2014 to present

The text was scraped from the Federal Reserve web site using Beautiful Soup and then vectorized using nltk and sklearn.



ARIMA and ARIMAX models were fit to the changes in interest rates for each date in the cross-validation dataset using pyflux and a one period forecast of changes in rates was calculated for each day. The time series models were also fit on the principal components of the entire set of interest rates using sklearn and fit to ARIMA and ARIMAX models.



Forecasts for the changes in interest rates over the next day for each date in the cross-validation set were compared to actual changes in interest rates to determine if the time series models provide a different distribution of rates.

Results:

The Federal Reserve speech text provides very little improvement forecasting changes in interest rates. While

I created an estimate of how different the text from each press release and speech was from the most recent texts, I did not compare this to the market's expectations given the same economic data. Finding variables that reflect the market's expectations may be an area for future study.

Technology used:

Beautiful Soup • Python • Numpy/Pandas • Scikit-Learn • Matplotlib • PyFlux • nltk

Github: [davidjsmith44](#)