Data 620 Final Project Proposal

By Alexander Ng and Philip Tanofsky

Date: April 26, 2022

Topic: Text analysis of FOMC Meeting Statements

Background: The Federal Reserve FOMC meetings produce policy statements that are closely studied by financial markets, politicians and business leaders.   The statements announce target Fed Funds rates, describe the Central Bank's views on economic conditions and asset purchases.   Substantial text analyses of these statements has been undertaken by machine learning researchers.

Project Objectives:

By analyzing the text corpus of FOMC policy statements between 2007-2022, we seek to understand how they have changed in response to and in advance of concurrent economic conditions using Latent Dirichlet allocation (topic models) and, if time permits, also with sentiment analysis (time series regression).   The latter would involve lead-lag regressions between sentiment (and their changes) to financial returns.

Data Sources:

FOMC Statements from 2007-2022

Economic time series from 2007-2022 such as real GDP growth, core CPI inflation, unemployment rates.

Financial time series from 2007-2022 such as a broad US stock index, US Treasury bond yield, IG credit spread subject to data availability from public sources.

Sentiment dictionary

Methods

We will use the NLTK library for textual preprocessing, scikit-learn for LDA, seaborn or other Python libraries for data visualization.

Roles and Responsibilities

Alexander Ng - background research, time series collection, LDA analysis

Philip Tanofsky - textual analysis, sentiment analysis

Certain tasks will be done jointly: presentation and video, exploratory data analysis and discussion of results.