STAT410 - Modeling US Inflation Project Proposal

Benjamin Allen

10/20/2022

Overview

The primary aim of this project is to model inflation in the United States using a variety of economic indicators. Given the nature of this project, the data is split up into numerous different data sets, where essentially every variable comes from a different source. Every data set has the following columns: year, month, and the variable we care about.

Data Sources:

1. Inflation Rate and Unemployment Rate from U.S. Bureau of Labor Statistics
2. Federal Funds Rate, Raw Materials Price Index, and Consumer Sentiment Index from St. Louis FRED
3. Oil Price Index from U.S. Energy Information Administration

Model Variables

The variable of interest to model is the 12-month percentage change in the consumer price index which measures inflation experienced by consumers. This is equivalent to the inflation rate for consumer goods.

Traditional economics suggests that an inverse relationship exists between inflation and unemployment commonly known as the Phillips curve, so we will use monthly unemployment data as a regressor.

Another potential predictor of inflation is the target rates set by the United States Federal Reserve, as these are often changed to affect inflation.

Another factor believed to drive inflation is increases in production cost which could be modeled by using monthly market prices for different raw materials as regressors. For now, oil and general raw materials prices seem like good options.

Finally, it would be valuable to consider the demand side of inflation, which can be modeled using the consumer sentiment index developed by the University of Michigan.

Methodology

The first approach will be a multiple linear model, and a number of specifications have already been tested. From initial testing of these models, linearity and homoscedasticity is questionable but not obviously untrue. Next, the residuals are normally distributed and have been verified visually and with a shapiro test. Finally, with time series data the biggest concern for a linear model will be autocorrelation and after performing a Durbin Watson test for autocorrelation, all linear models tested show an autocorrelation > 0.7.

Given this result, it will be best to model this data using either an ARIMA model or an exponential smoothing model. Both approaches will be tested.