Forecasting Inflation with Machine Learning

Author: Sarah Larino

Advisor: Dr. Andrii Babii

Abstract

In this thesis, I employ a number of machine learning (ML) methods on the inflation forecasting problem space. I utilize macroeconomic indicators alongside textual data and apply ML methods to an updated time horizon. Ultimately, I find that ML methods are a viable alternative to traditional benchmarks under certain time horizon conditions, particularly with the inclusion of textual data. However, in contrast with the previous literature, I demonstrate that some ML models are particularly sensitive to the treatment of outliers. When a full time horizon is employed and outliers are included, certain ML models that performed well in previous analyses are not able to outperform other forecasting methods.

Data

The primary data source for my thesis is FRED-MD, 134 monthly U.S. indicators curated by the Federal Reserve Bank of St. Louis. The time series indicators included extend from January 1959 to August 2022.

I also introduce textual data sourced from Structure of News, a dataset crafted by financial economists at the Yale School of Management that summarizes Wall Street Journal articles from 1984 to 2017 into monthly topical themes. Ideally, this textual data serves as a proxy for inflation expectations and catches nuances that are not represented in more traditional data sources. When using this data, I employ a preselection step to only choose topics that demonstrate predictive power for inflation.

Methods

The Personal Consumption Expenditures Price Index (PCEPI) serves as my key response variable, inflation, in my model. PCEPI measures the prices of goods and services paid by consumers in the U.S. and is typically chained to a base year, 2012. PCEPI is the preferred inflation measure of the Federal Reserve and is released by the Bureau of Economic Analysis. I transform PCEPI to be stationary by taking the monthly log difference.
Additionally, when creating my forecasts, I employ a rolling window. This makes my forecasts more robust to structural changes.

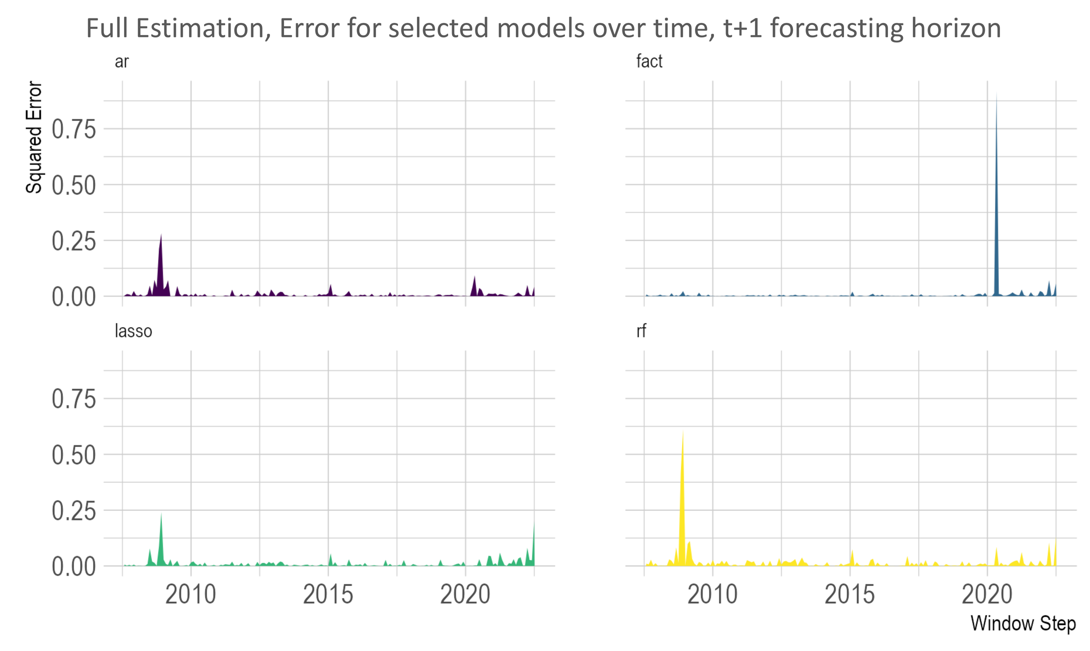
Contributions

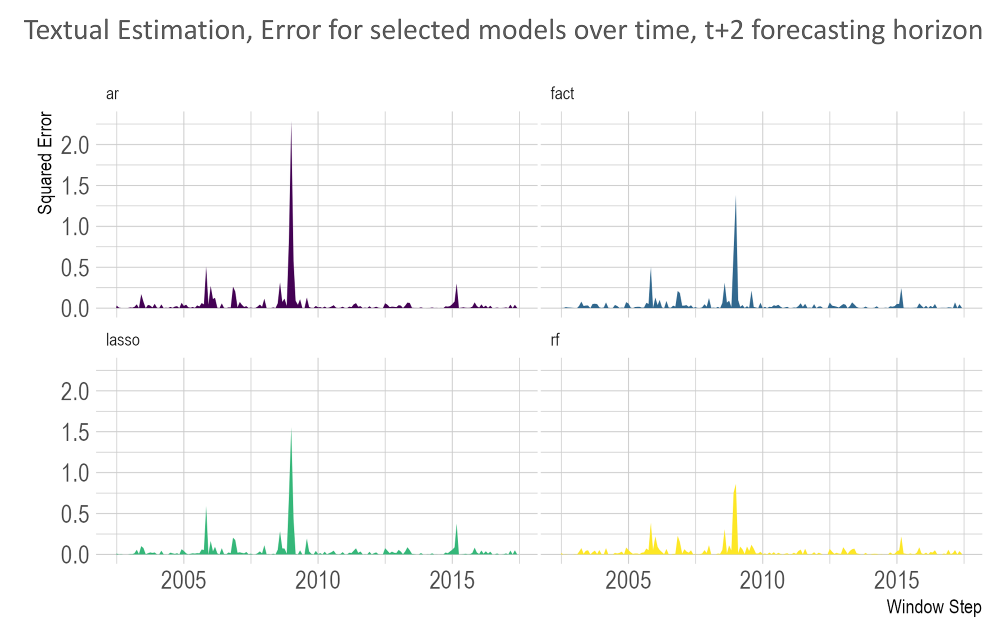
1. I employ **textual data**, which has not been previously utilized in inflation forecasting.
2. I apply ML methods to an **updated time horizon** that encapsulates both the unusual low inflationary period that occurred between 2008 and 2021, and the period of unusually high inflation that began in 2021 and has persisted to the present. (Previous research was limited to a forecast horizon that ended in January of 2016.)
3. I **do not remove extreme data points** during my estimation process. Outlying data was removed in the previous literature, particularly points realized during the 2008 Global Financial Crisis, to aid the performance of the forecast models.
4. I **modify the implementation of lags.** When fitting my LASSO regression model, this reduces poor performance caused by over-penalization. I parameterize lags according to Legendre polynomials in order to reduce the level of dimensionality present.

Models

I estimate three benchmark models (*italicized*). These models encompass both the classic univariate time series models and the theoretical-based models of inflation. I then estimate a number of models that will be compared with the benchmarks. The ML models that are particularly salient are LASSO and RF.

Random Walk	*RW*	*Naïve; uses current inflation*
Autoregressive	*AR*	*Direct; # of AR terms determined by BIC; parameters estimated via OLS*
Phillips Curve Motivated	*PC*	*Based on lagged inflation and unemployment rate*
Factor	Fact	Latent factors extracted and regressed on to produce forecast
Bagging	Bag	Aggregates OLS models across bootstrap samples
Complete Subset Regression	CSR	Averages across all possible models after preselection step
Least absolute shrinkage and selection operator	LASSO	Performs regularization on OLS to penalize the abundance of predictors; creates sparse model
Random Forest	RF	Creates decision trees from bootstrap samples, then averages over all trees
Combinations	Comb	Combines forecasts to 'hedge' their errors





Results

In the full estimation, I find that the best performing model for shorter and very long horizons is the factor model. For the mid-term forecasts, the AR model outperforms all others. However, during the 2008 financial crisis, both the factor and LASSO models are better than the AR model at handling the shock, realizing a smaller spike in errors during this period. During the Covid-19 pandemic, the one month ahead factor forecast sees a monumental spike in forecasting error, while the other models handle the shock with relative ease.
In the textual estimation, limited to the time period between 1984 and 2017 because of the available data, the RF model is able to outperform the AR model, especially in longer forecast horizons. The factor model is only the best performing for the one month ahead forecast. Both the LASSO and RF models perform particularly well during the 2008 crisis.
In examining the variable selection for the LASSO and RF models, I see that the level of model complexity tends to be less for very short term forecasts. This complexity increases as the forecasting horizon increases, before eventually becoming more simplistic again for the longer term horizons. This excitingly aligns with many of the theories of inflation drivers.

Conclusions

I demonstrate that **the specificities of ML models are incredibly important to their ultimate efficacy as forecasting tools**. When including the outlying data that previous literature did not, ML models are not able to outperform other forecasting methods, with the RF model appearing to be the most sensitive to these outlying data points. Similarly, the treatment of lags can significantly affect ML model performance. In the case of the LASSO model, by altering the treatment of lags from that which was utilized in previous literature, I am able to improve the model's performance. Via the inclusion of textual data, I am able to improve the performance of ML models, particularly the RF model, such that they outperform standard benchmarks in most forecasting horizons. However, this estimation is limited by the time horizon of my available dataset.

ML models in my research do show promise in their usefulness as forecasting tools. They **accommodate well for nonlinearities** and do not rely as heavily on underlying assumptions about theoretical inflation drivers, making them **more flexible during times of unusual economic activity**. This ability to cope with economic crisis could be useful for the policy making of central banks during periods when other forecasting models flounder. Both the LASSO and RF models also showed stability in their variable selection mechanisms.
Notably, there appears to be **promise in the incorporation of textual data**. This is an area of inflation forecasting that is novel to my research and should definitely receive further attention and testing going forward, especially as this type of data becomes more readily accessible. Ultimately, **further research on inflation forecasting should consider ML models**, but more attention needs to be given to the tuning of these ML tools in order to better understand how they could be employed.

Contact: Sarah Larino
sglarino@gmail.com