Jude Mersinger

U21950507

ECO 4422

May 1, 2024

Final Project Write Up

Data Selection

The data used in this project come from the FRED economic database. The variables to be analyzed are Median Consumer Price Index (CPI), Trade Balance, and Unemployment Rate, all for the US. The time range for analysis will be from January 2000 to December 2019.

Initial Analysis

The first step I took in this analysis was to graph all three variables over time. In analyzing the graphs, all three variables appeared to be exhibiting covariance stationarity. This observation was reinforced in the graphs of the autocorrelations and partial autocorrelations of each variable, which showed the autocorrelations approaching 0 along τ for CPI, Trade Balance, and Unemployment.

Before starting to build the VAR model, I chose to check the cross correlations to ensure the data has some relationship (cross correlation values are lines 30-74 in the code). This analysis showed nonzero cross correlations, so I chose to move forward to creating the VAR model.

Vector Autoregression (VAR) Model Construction

To select the lag for my VAR model, I used the @varlagselect command in RATS, looking at AIC and SBC for information criteria. By AIC, a lag length of 7 was best for the model. By SBC, a lag length of 2 was best for the model. I chose to go with the lag length suggested by SBC, as SBC is more consistent and parsimonious.

After running the VAR(2) model, I observed that Unemployment exhibits predictive causality on CPI and Trade Balance, but not vice-versa. This means that Unemployment exhibits Granger causality on both CPI and Trade Balance. Neither CPI nor Trade Balance exhibited predictive causality on other variables. This will be important for the final comparison and analysis for the VAR and ARIMA models.

After analyzing the VAR(2) model, I ran a truncated model to December 2018, and used this model to forecast CPI, Trade Balance, and Unemployment from January 2019 to December 2019. I will look at MSE for the VAR model at the end, so I moved on to building the ARIMA model.

Autoregressive Integrated Moving Average (ARIMA) Model Construction

To start constructing my ARIMA models for the data, I ran the @bjautofit command to find the best ARMA(p,q) model. I used SBC as my information criteria and set pmax = qmax = 5. The optimal ARMA(p,q) models came out to the following: ARMA(1,1) for CPI, ARMA(4,0) for Trade Balance, and ARMA(2,2) for Unemployment.

I then ran a Unit Root Test to determine if any of the variables require differencing. In the test, the null hypothesis was the existence of a unit root. Of my three variables, only CPI rejected the null. This meant that Trade Balance and Unemployment had evidence of a Unit Root.

To ensure I had the best fitting models for these variables, I decided to run models for Trade Balance and Unemployment with and without differencing, ARMA(4,0) and ARIMA(4,1,0) for Trade Balance, and ARMA(2,2) and ARIMA(2,1,2) for Unemployment. The information criteria (both AIC and SBC) for these models suggested that the ARMA models for both variables was better than ARIMA. I decided to move forward with the ARMA models to forecast.

Like for the VAR(2) model, I truncated the models to December 2018, and forecasted each variable from January 2019 to December 2019. Once both VAR and ARMA models have been forecasted for each variable, I was able to compare them.

VAR vs ARMA/ARIMA Model Comparison

I started my comparison by graphing the forecasts alongside the actual values from January to December of 2019. For both CPI and Trade Balance, the VAR(2) model appears to be closer to the actual values. For Unemployment, both models appeared to be very similar, so an analysis of MSE would be necessary.

After running @uforeerrors for each model, I observed that the MSE was smaller in the VAR model for all three variables. However, it should be noted that the MSE for Unemployment only differed by about 0.03.

Conclusions

The MSE analysis shows that the VAR model was better for predicting this data. These MSE values match what was observed when I looked at the predictive causalities. Before making the VAR, I observed that Unemployment exhibited Granger Causality on CPI and Unemployment. This means that including Unemployment in the VAR model improved the forecast of CPI and Unemployment compared to CPI or Unemployment by themselves (as in the ARMA model). The difference in MSE is so small for Unemployment because there was no variable exhibiting predictive causality on Unemployment to improve its forecast.