Michael Cao (yc849), Winston Park (wp254)

Time Series Analysis of United States Labor Force Participation Rate

Introduction

In economics, researchers use the Solow Growth Model to project the aggregate GDP of the US economy as a function of capital, technology, and labor. We are particularly interested in studying the labor force participation rate (LFPR) to examine the third key factor that drives economic output. A higher LFPR indicates that more individuals in the country are producing goods and services, leading to higher output levels. Indeed, if we analyze the prosperous period following World War II, there exists a clear correlation between a growing LFPR and GDP growth[1]. We forecasted LFPR to help understand the future direction of the broader economy.

We gathered non-seasonally adjusted datasets for the following metrics: Labor Force Participation Rate, Consumer Price Index, Unemployment Rate, and Population Level. All data was collected by the US Bureau of Labor Statistics and retrieved from the Federal Reserve Bank of St. Louis. We used monthly data from January 1948 to December 2019 because the economic shutdown following the onset of the COVID-19 pandemic presented large outliers that make forecasting difficult. LFPR measures the proportion of individuals in the population who are 16 years or older that are in the labor force. This figure gives economists the data necessary to assess labor market health. Consumer Price Index (CPI) is a measure of the price levels of goods and services in the economy. A change in the CPI value indicates a change in price levels in the economy (inflation/deflation). The Population Level dataset allows us to measure the overall size of the population that would be eligible to be included in the labor force. Finally, the unemployment rate is measured to understand the proportion of individuals in the labor force who are unemployed.

To prepare our data, we took differences in the LFPR dataset to account for inherent seasonality - LFPR seems to reach a low in January and peak around July. Consequently, we differenced the LFPR data and used a lag of 12 to mitigate the seasonality effects. Furthermore, we found that taking first differences on the LFPR dataset allows it to pass the Augmented Dickey-Fuller Test (a test for stationarity). Accordingly, we decided that it is best to use differenced data to forecast LFPR.

In addition to forecasting the labor force participation rate to build our understanding of the future direction of GDP, we would also like to use our time series models to investigate underlying relationships between LFPR and other economic phenomena. Our first model forecasts LFPR with a seasonal autoregressive integrated moving average (SARIMA) + generalized autoregressive conditional heteroscedasticity (GARCH) model to explore the potential economic cycles. Our second model forecasted LFPR through the vector autoregressive with exogenous variables (VARX) model, which is constructed to understand the relationship strength between LFPR and other key metrics like inflation and population level.

Through the forecasts provided by both models, we are able to understand the short-term direction of LFPR. We can apply this prediction, along with information on capital, technology, and other labor inputs, to build our understanding of the future path of US economic output. We hope to bring insights into labor productivity in the US, which is a key factor in macroeconomic

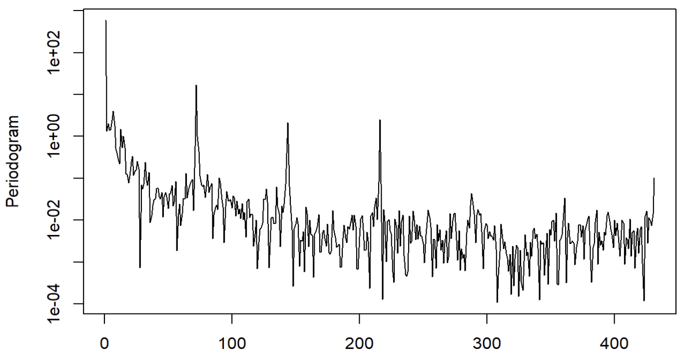
[1] The graphs for US GDP and LFPR from 1948 - 2000 can be found in Appendix I.

models used to calculate and predict GDP. As such, this paper would be of most interest to macroeconomic researchers, policy makers, and financial analysts who seek to enhance their understanding of the drivers of the American economy.

Results

i) Model 1: Seasonal Autoregressive Integrated Moving Average (SARIMA) + Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model

We first employed a SARIMA + GARCH model to forecast LFPR. To start, we found the periodogram for our LFPR dataset, which is shown below:



The periodogram peaked at frequencies of 72, 144, and 216, which are all multiples of 12. This pattern may indicate the presence of harmonics in the data that are occurring due to underlying economic cycles of 12 months, 6 months, and 4 months. Hence, we considered seasonally differencing the data using S=12 as our seasonality component for our SARIMA model, since 1/12 is the main frequency.

We plotted the autocovariance function (ACF) and partial autocovariance function (PACF) of our LFPR dataset and found that the PACF spikes after 1 year[2]. We checked the ACF and PACF of the seasonally differenced dataset to confirm this spike[3], and we concluded that we would use S=12 as the seasonality component for our SARIMA model.

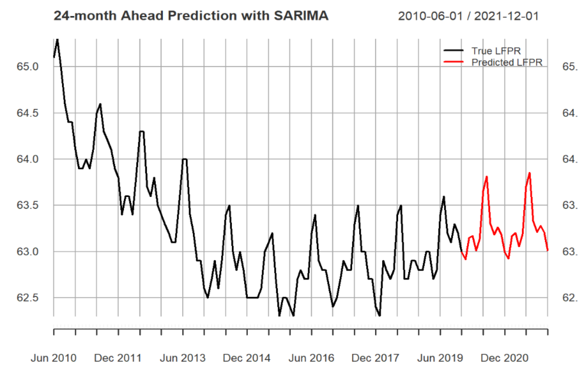
To fit our SARIMA model, we used the auto.arima function in R to find the best fit for our ARIMA parameters. This method returns ARIMA parameters of p = 2, d = 1, and q = 2, which gives us the lowest Akaike's Information Criteria (AIC) value. After studying the ACF and PACF of the seasonally differenced data, we determined the seasonal autoregressive order (P) to be equal to 1 and the seasonal moving average order (Q) to be equal to 2 because the periodic trend does not persist. Thus, we fit a SARIMA $(2,1,2)(1,1,2)_{12}$ to LFPR.

We then examined the diagnostic tests for our SARIMA $(2,1,2)(1,1,2)_{12}$ model[4]. Although the ACF plot and QQ-plot of the residuals both look relatively stable, confirming part of the assumptions, the standardized residual plot showed significant signs of heteroscedasticity - the variability of the residuals decreases when lag values increase, and the Ljung-Box statistic has p-values very close to 0, indicating errors that are not independent. We decided to fix this by fitting a GARCH model on the residuals of the SARIMA model.

In general, lower ARCH and GARCH orders tend to be better fits for data than higher orders when the data exhibits low evidence of volatility clustering. When observing the residuals of the model[5], we saw that they demonstrate heteroscedasticity but are not particularly clustered.

[2] The ACF and PACF graphs for LFPR can be found in Appendix II.
[3] The ACF and PACF graphs for differenced LFPR can be found in Appendix III.
[4] The diagnostic results for the SARIMA $(2,1,2)(1,1,2)_{12}$ can be found in Appendix IV.
[5] A plot of the residuals can be found in Appendix V.

Here, a GARCH order of 1 and ARCH order of 2 returned a p-value much higher than 0.05 for different lags on both the weighted Ljung-Box test of standardized residuals and standardized squared residuals. Although the Pearson goodness-of-fit test did not return p-values greater than 0.05 for some lags, it solved the issue of heteroscedasticity in our original SARIMA residuals.



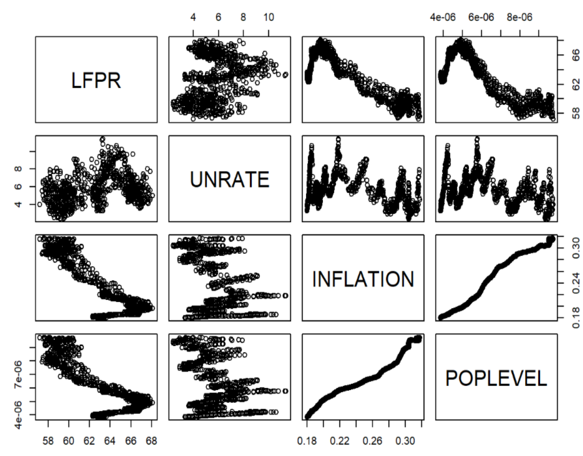
We made a 24 month ahead (2 year) prediction with the SARIMA $(2,1,2)(1,1,2)_{12}$ model using the sarima.for function in R. The graph for this short-term prediction can be found above (forecasted data in red): The SARIMA + GARCH model predicted that the LFPR will remain relatively stable from 2020-2022 with a slightly positive underlying trend.

ii) Model 2: Vector Autoregressive Model with Exogenous Variables (VARX)

We would like to go one step deeper in our analysis to investigate if there are any other key economic measures that would have significant influence on LFPR and consequently improve our model fit, while still taking the periodic behavior into account. To capture the dynamic relationships between several variables over time, we employed the vector autoregressive model (VAR), a model widely used in macroeconomic analysis to assess the effects of policy interventions or shocks on the economy.

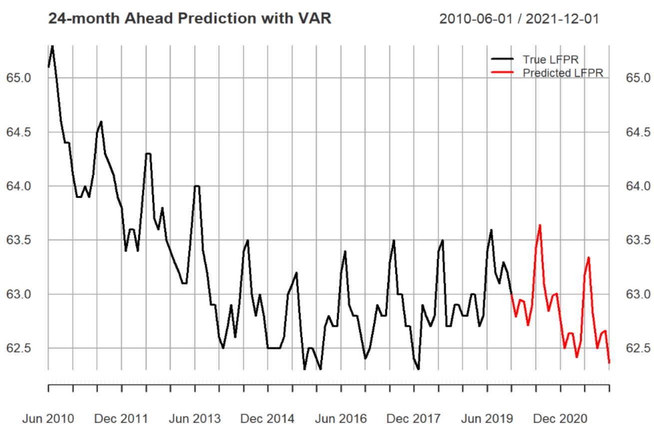
We initially considered three variables to be included in the VAR model in addition to LFPR - unemployment rate, consumer price index, and population level. Before training the VAR model with these variables directly, we needed to analyze whether doing this is reasonable. To begin, we computed the correlation matrix and plotted the scatterplot matrix[6], where none of the pairs seemed to have a strong linear relationship with LFPR; CPI and population level seemed to be non-linearly correlated with LFPR, whereas the unemployment rate, to our surprise, does not seem to have any correlation. We then attempted transformations on these variables, and ultimately, we were able to observe very strong pairwise correlations between LFPR, the inverse-log of CPI, and the inverse of population level, as shown in the correlation matrix and scatterplot matrix on the right. The R-squared values of simple linear regressions regressing each variable on LFPR also provided another layer of confirmation[7]. Consequently, we decided to exclude unemployment rate, and train a VAR model with three correlated variables: LFPR, the inverse-log of CPI, and the inverse of population level.

	LFPR	UNRATE	INFLATION	POPLEVEL
LFPR	1.0000000	0.2547739	-0.8712033	-0.8247537
UNRATE	0.2547739	1.0000000	-0.3161169	-0.3087008
INFLATION	-0.8712033	-0.3161169	1.0000000	0.9893472



[6] Here, 'INFLATION' denotes the inverse-log of CPI, and 'POPLEVEL' denotes the inverse of population level.
[7] The summary output of the linear regression results can be found in Appendix V.

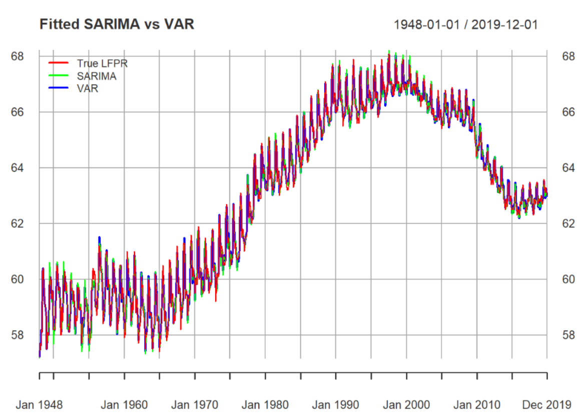
We considered fitting a VARX model using different frequencies as exogenous variables as well as a simple VAR model with no exogenous variables. The frequencies considered included all the combinations of the ones detected previously by the periodogram (1/12, 1/6, and 1/4) expressed as a combination of sine and cosine functions, and AIC values were computed for every model. Ultimately, we found the lowest AIC model is a VARX model where the exogenous variables consist of a trend component and a combination of all three frequency components. This model has an AR order of 25. This model is a good fit of the LFPR data, as the correlation matrix for the residuals resembles a diagonal matrix, with non-diagonal terms very close to 0, indicating that the linear dependencies and patterns in the data are successfully captured[8]. The plot on the left shows the LFPR forecast for the next 24 months using this model.



In addition to obtaining a VARX model, we would also like to investigate the significance of the causal effects of each variable pairs with Granger causality, a tool commonly used in economics to analyze the interdependencies between economic variables. Using the same exogenous variables in our model, we yield a p-value of 0.003235 for the null hypothesis that the inverse-log of CPI do not Granger-cause LFPR, and a p-value of 0.07242 for the null hypothesis that the inverse population level do not Granger-cause LFPR. Therefore, the inverse-log of CPI seems to be a more significant factor in predicting the changes in LFPR; and although the inverse population level fails to reject its null hypothesis at the 0.05 significance level, it is still strongly correlated with LFPR and arguably still a reasonable variable to be included in our model that can potentially improve its fit.

Model Comparison

The two models used in this report were created to answer different questions about LFPR. In particular, the SARIMA model serves to identify short term cycles in the US economy, while the VARX model pinpoints key economic metrics that have a significant impact on LFPR. Comparing the fitted lines[9], we observe that both models are able to achieve a close fit.

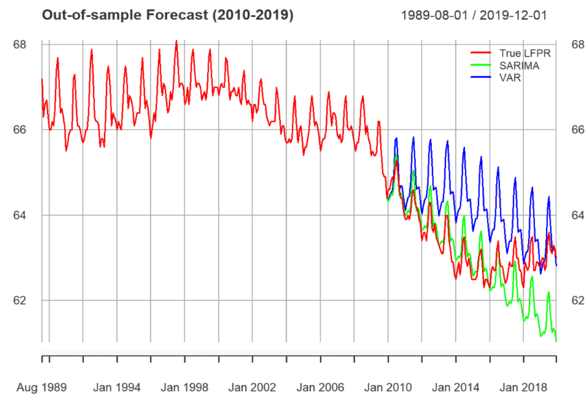


The performance difference of models, however, is amplified by the mean squared forecast error. This has been computed and visualized by using the out-of-sample method, where the data from 1948 to 2009 are held as the training data: a 120-month forecast was generated using both models, then compared against the LFPR from

[8] The correlation matrix of the VARX model residuals can be found in Appendix VI.

[9] The fitted models focused on separate time periods can be found in Appendix VII.

2010 to 2019, which is subsetted as the test data. The mean squared forecast error of the VAR model is 1.209, while for the SARIMA model is 0.516, which indicates a better forecast ability. However, this should not be taken at its face value, as a different cutoff time point as the train-test split would quite reasonably produce a different result, especially if that time point is a turning point in the economy, which is always highly volatile. Additionally, as shown in the graph to the right, although both models are able to forecast the short term (within 24 months) LFPR relatively well, neither one is able to capture the long-term trend successfully, due to unpredictive nature of the economy.



Achieving an accurate long-term forecast, however, is not the purpose of either model. The VAR model incorporates external factors that might impact our variable of interest, and is consequently a good model to identify the causality of these factors. On the other hand, although the SARIMA model does not incorporate any information beyond time, it is a suitable model to explore underlying economic cycles with its more sophisticated seasonality structure.

Discussion

In this study, both SARIMA and VARX models showed that LFPR is a relatively stable input for labor productivity. Moreover, through Granger Causality, we explored the relationship between LFPR and CPI and population levels. We concluded that the inverse-log of CPI Granger causes LFPR at a 5% significance level and the inverse population level, while weaker than inverse-log CPI, still Granger causes LFPR at the 10% significance level. Thus, we view these variables as inputs that should be included in models to predict LFPR. We predict that, barring the catastrophic effects of COVID-19, LFPR would have remained relatively stable from 2020-2022. Indeed, if we view the actual LFPR after June 2020 (after the large shock to the economy), we can see that LFPR returns to its normal patterns with a slightly upwards trend due mean reversion after the huge trough experienced from March to May 2020[10].

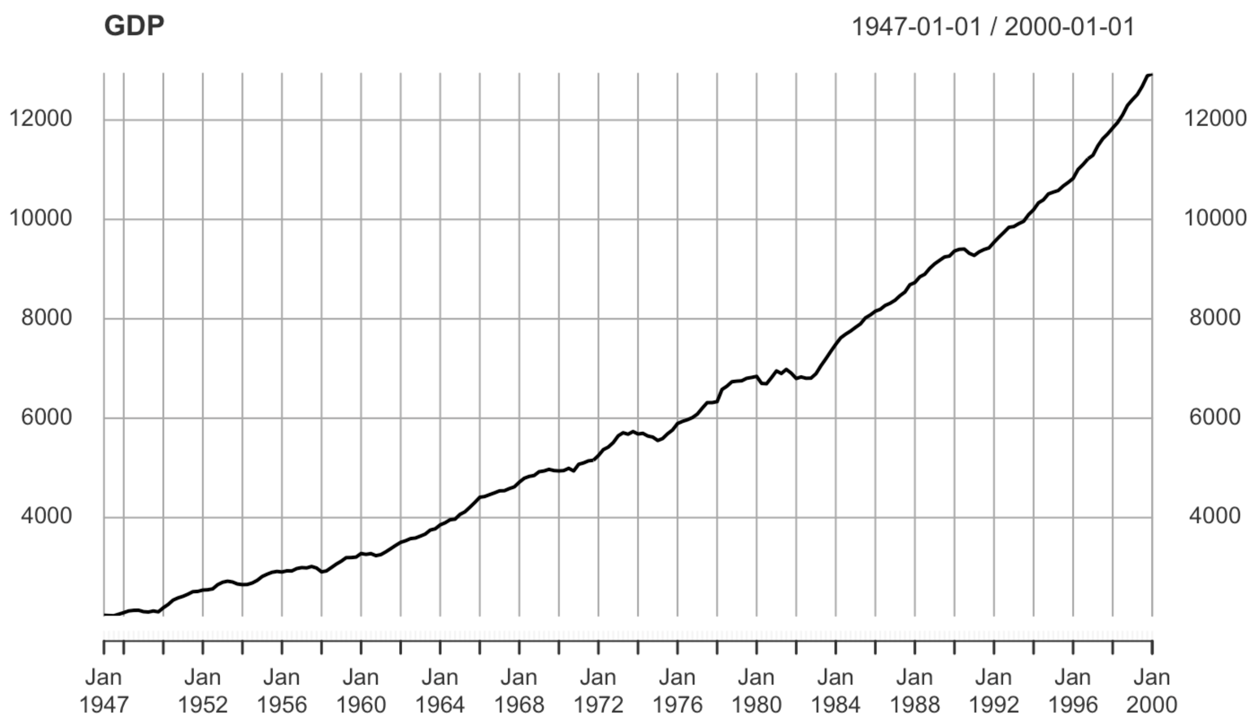
LFPR appears to follow broader economic cycles, which should prove useful when used in combination with other economic factors (e.g. unemployment, inflation, interest rates, etc.) to predict the future short-run path of the economy. SARIMA and VARX do a sufficient job at predicting the short-run path of LFPR. However, our models would not be particularly strong in terms of predicting long-run LFPR because they fail to account for longer term cycles in the data. Therefore, we recommend further study on this topic using a more sophisticated model to approach this problem. An exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS) model, which incorporates longer term cycles in the trend to help make predictions could provide a more reliable long-term forecast. We took initial steps into a TBATS prediction[11], which demonstrated superior fit to LFPR than the other models [12], but additional study into this method is needed.

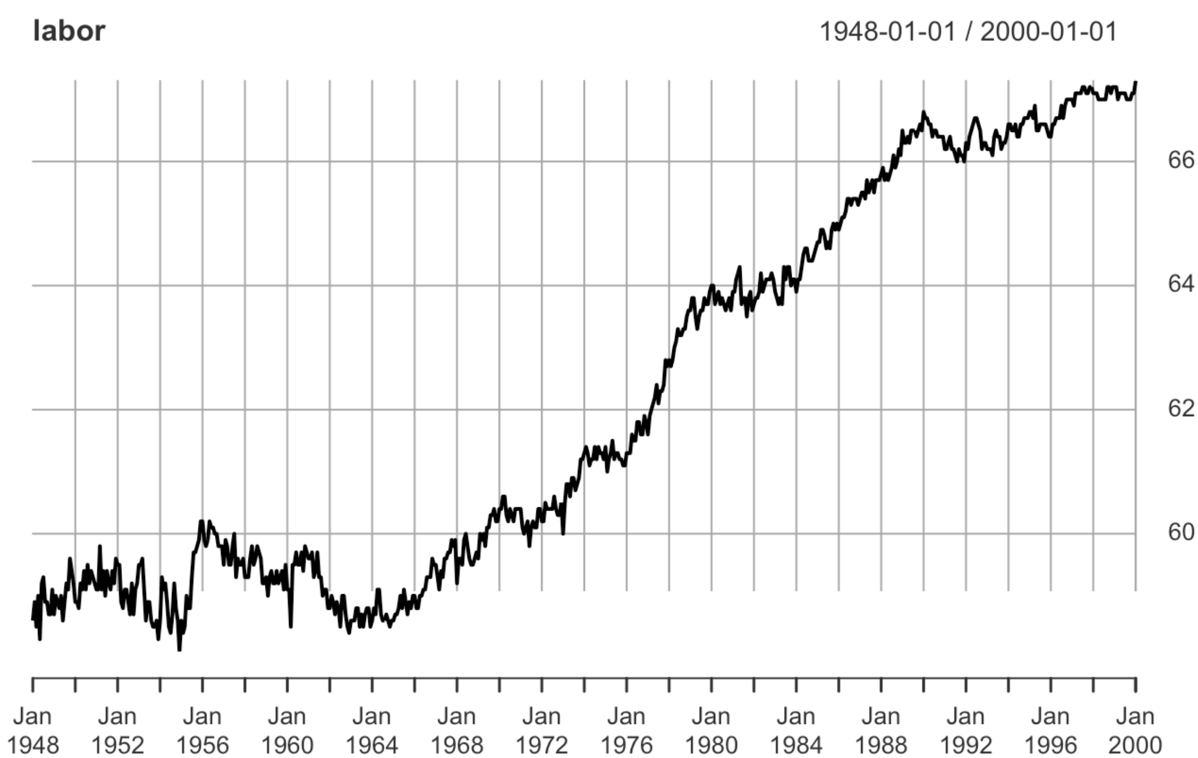
---

[10] A graph of the LFPR from June 2020 - April 2023 can be found in Appendix VIII.
[11] Our 24 month-ahead prediction with TBATS can be found in Appendix IX.
[12] The model fit for our TBATS model can be found in Appendix X.

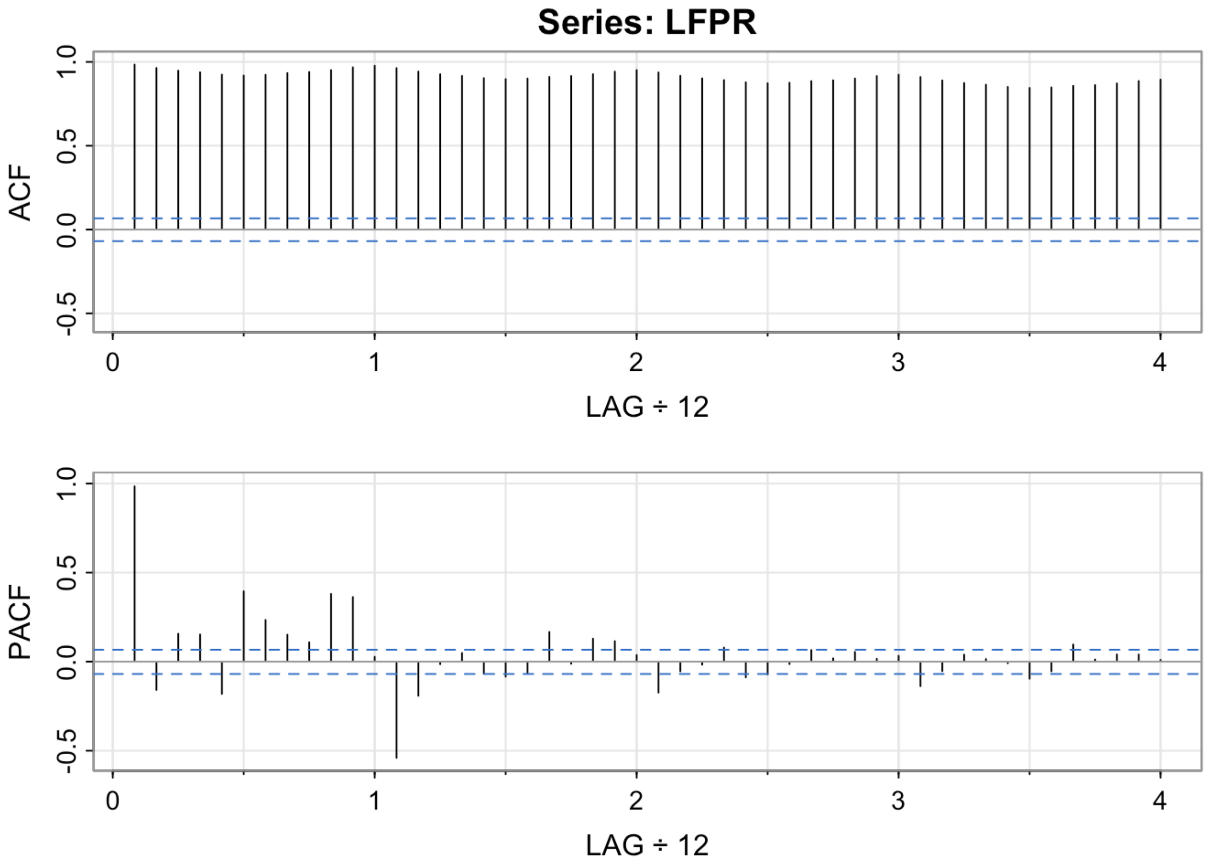
Appendix

I. <u>US Labor Force Participation (1948 - 2000) and Real GDP (1948 - 2000)</u>

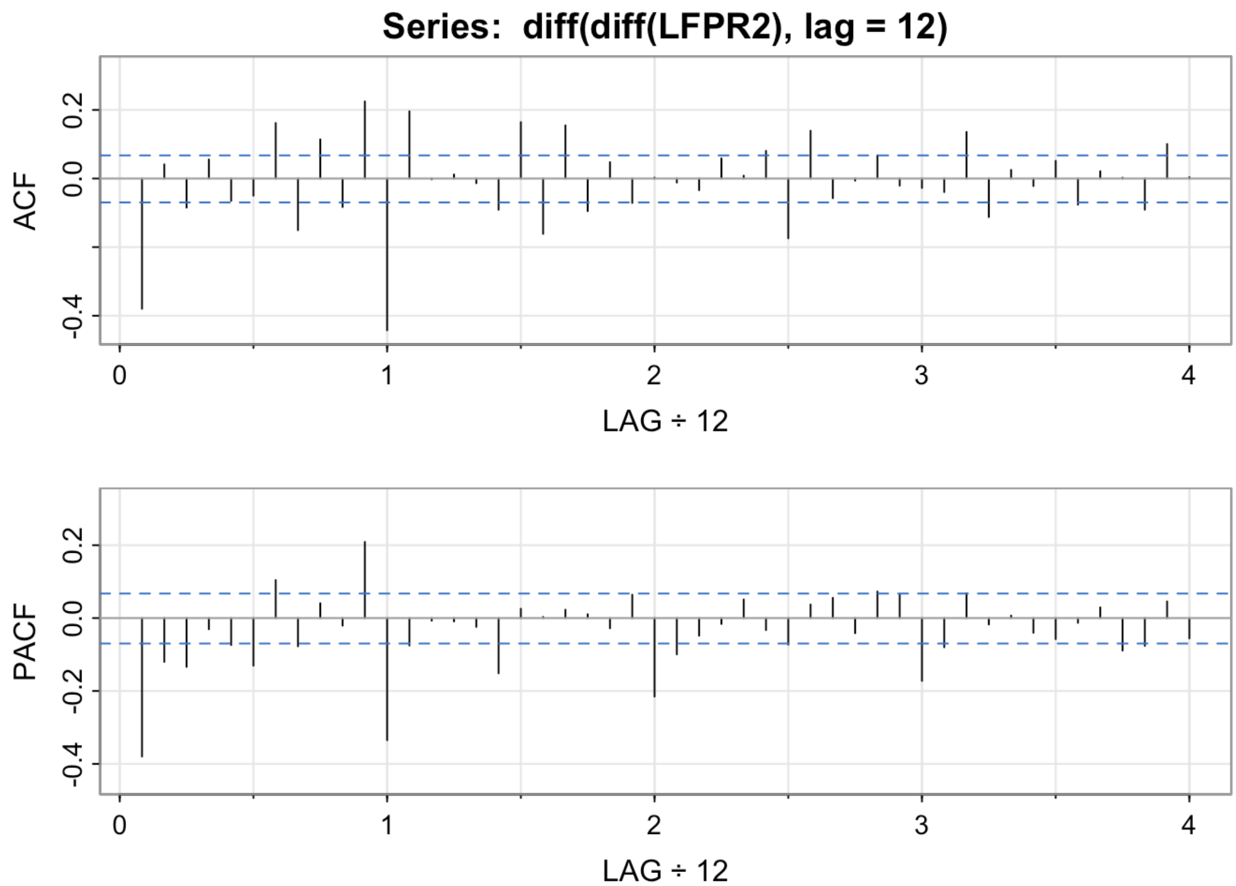




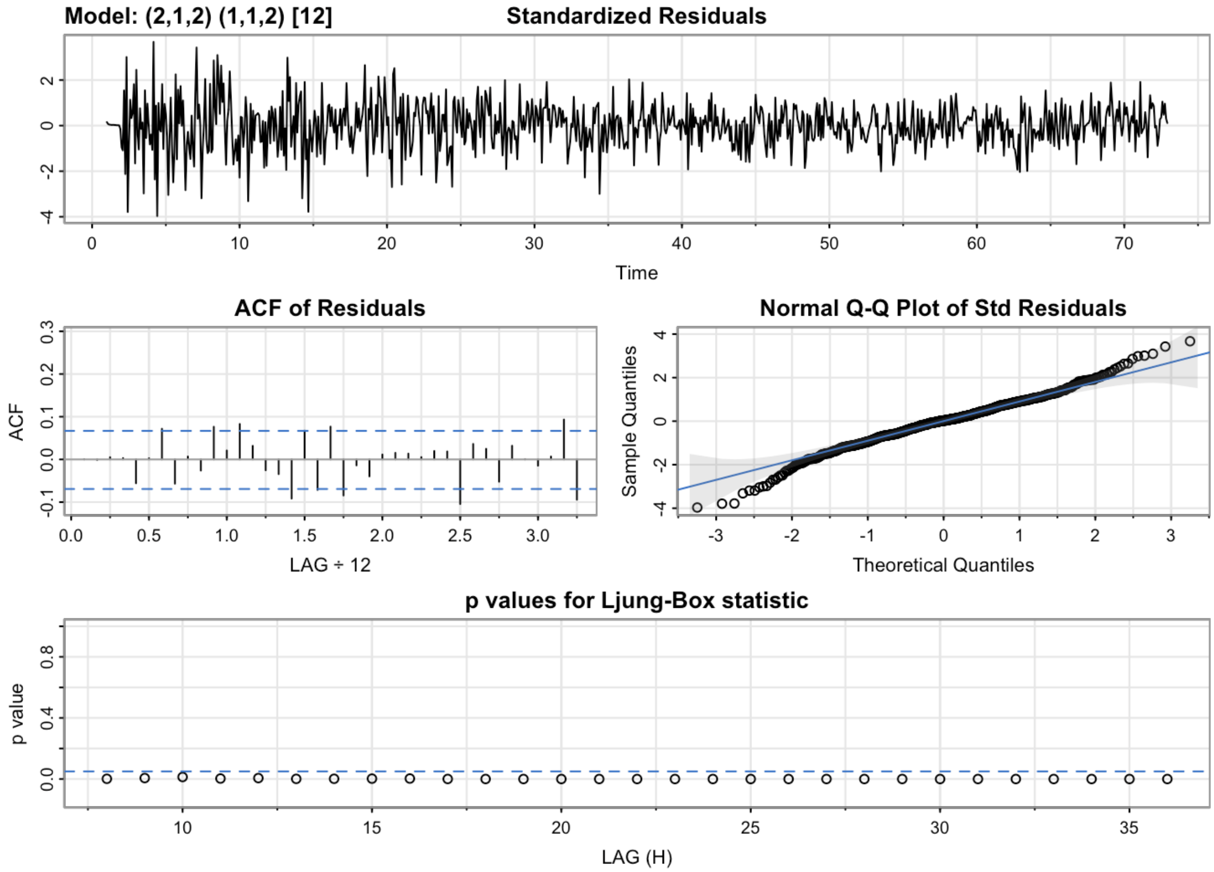
II. <u>ACF and PACF of LFPR</u>



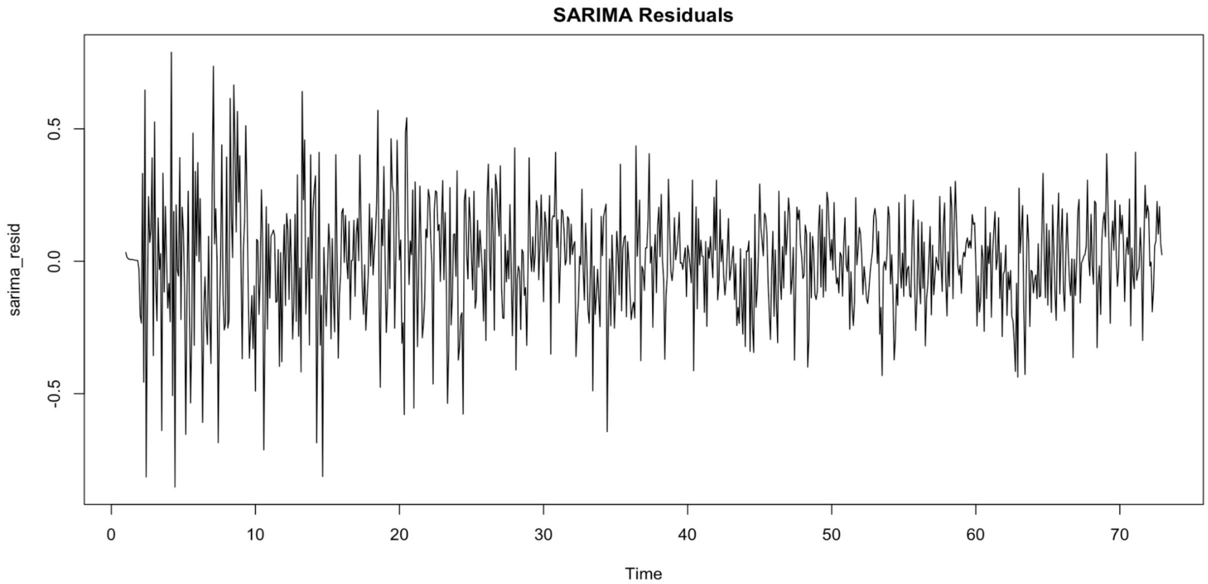
III. ACF and PACF of Seasonally Differenced Data



IV. SARIMA $(2,1,2)(1,1,2)_{12}$ Diagnostics



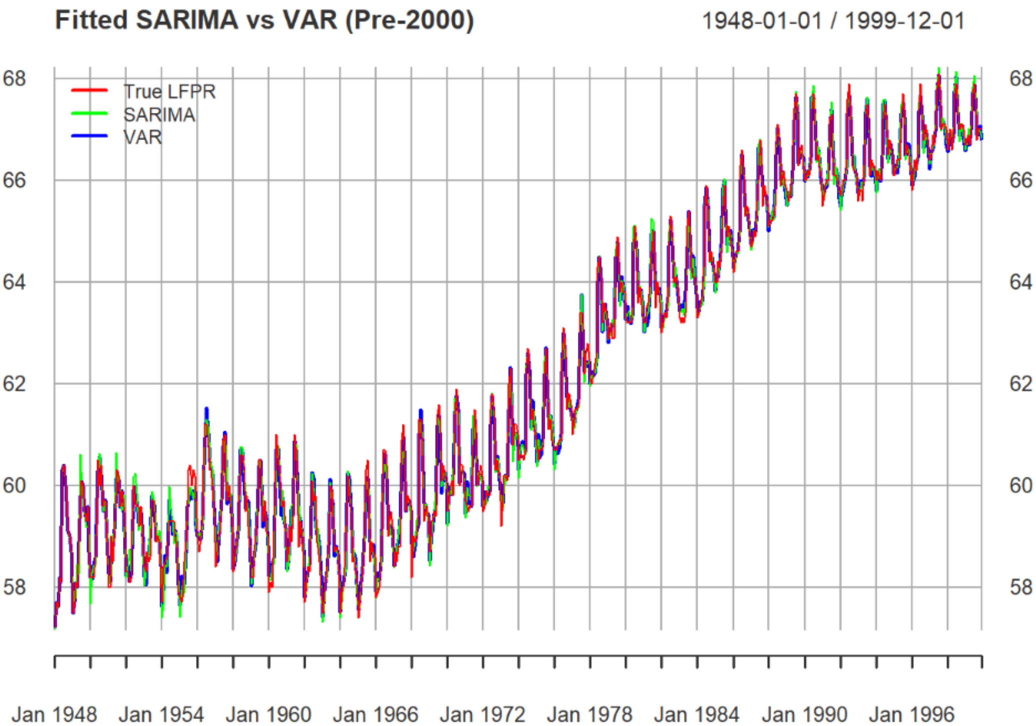
V. SARIMA Residuals Plot

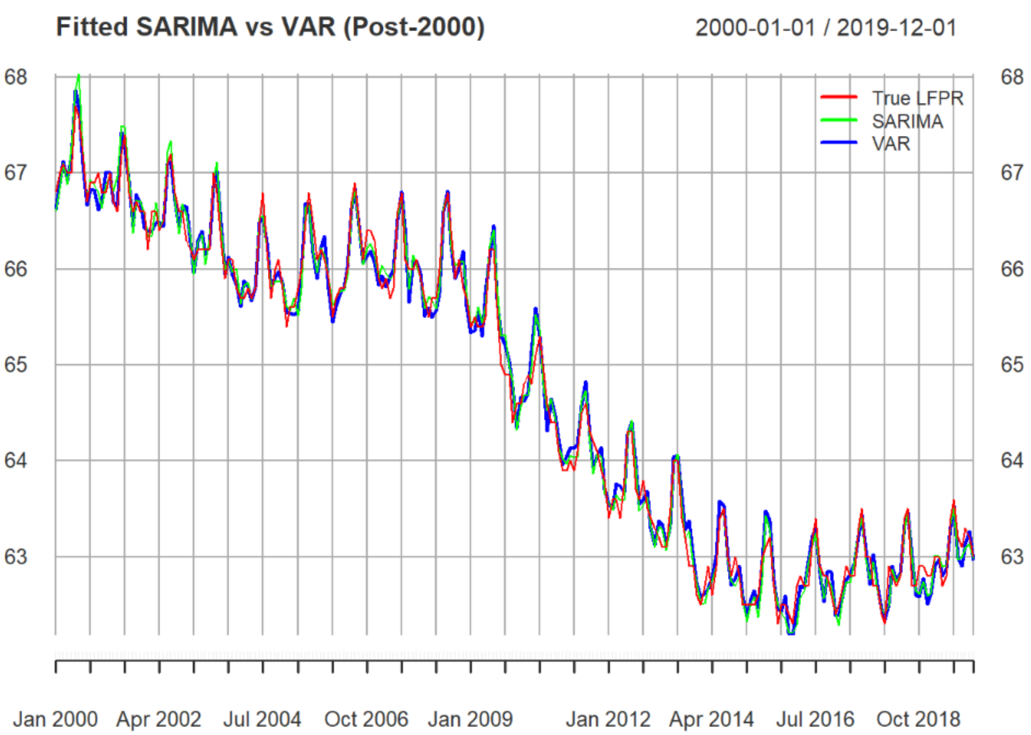


VI. Correlation Matrix for the VARX Model Residuals

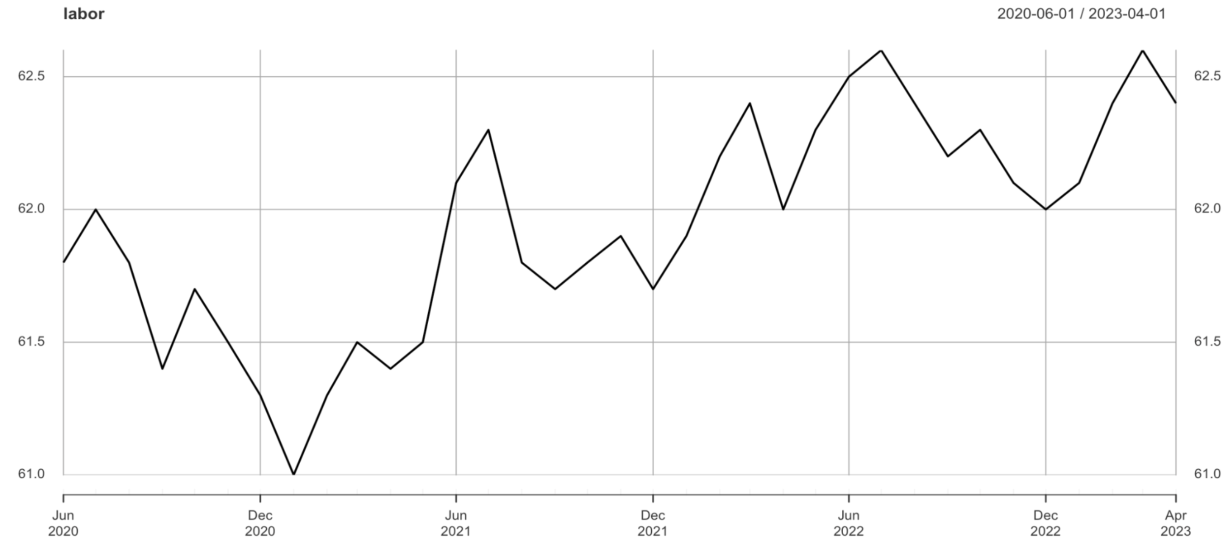
	LFPR	INFLATION	POPLEVEL
LFPR	1.000000000	0.06819137	-0.009944367
INFLATION	0.068191374	1.00000000	-0.086887176
POPLEVEL	-0.009944367	-0.08688718	1.000000000

VII. Fitted SARIMA and VAR Models Focused On Separate Time Periods

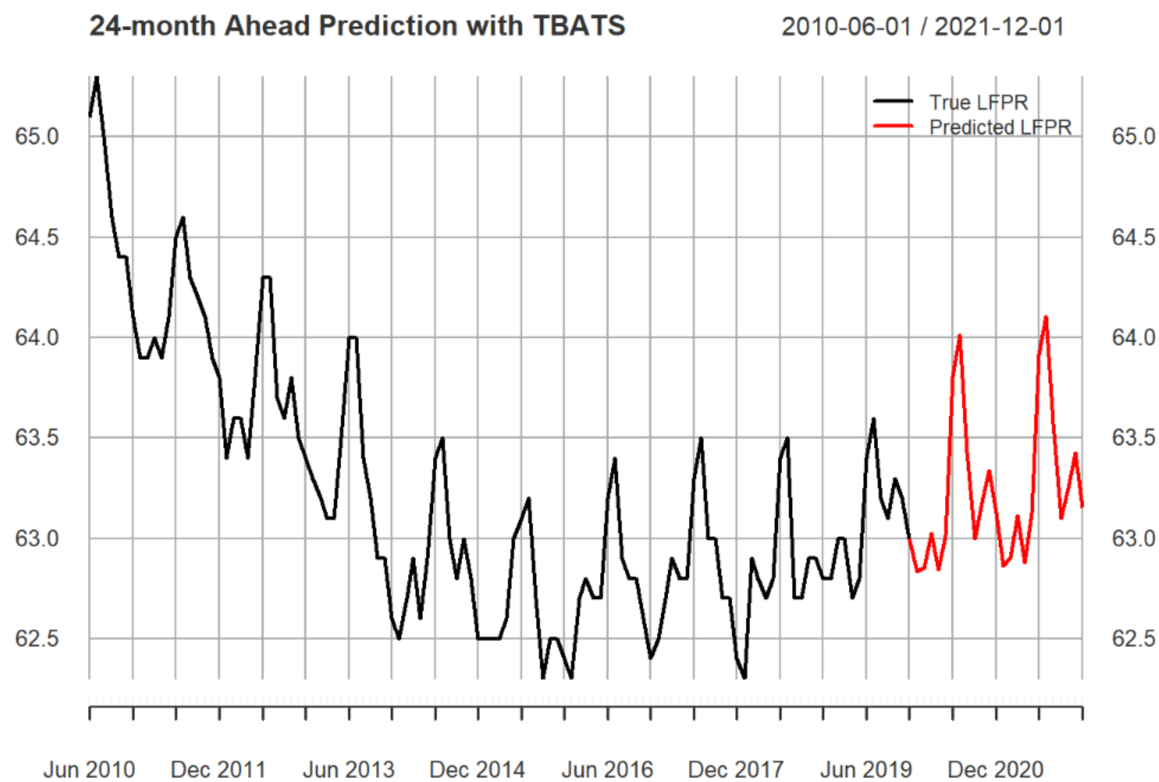




VIII. LFPR from June 2020 - April 2023



IX. 24-month Ahead Prediction with TBATS



X. Out-of-Sample Predictions with TBATS Model