ANLY 565: Time Series and Forecasting

Forecasting Inflation and Unemployment in the US Economy

Project Report

Subhash Pemmaraju
8-7-2020

INTRODUCTION

The ability to accurately forecast inflation is critical to planners across the world, particularly among central banks such as The Federal Reserve. One important foundation for how central banks conduct monetary policy is the relationship between unemployment and inflation as illustrated by the Phillips Curve[1]. According to the Phillips curve, there is a downward sloping relationship between inflation and unemployment. During periods of low unemployment, inflation picks up and during periods of high unemployment, inflation drops.

Over the last couple of years, there has been a lot of debate on whether the Phillips curve relationship continues to hold in the modern economy[2] from the late 1970s and much research has gone into why it doesn't hold as firmly as it did in prior to that period.

The aim of this paper is to explore the relationship between inflation and unemployment through time and to develop suitable time series models to forecast these variables.

MOTIVATION & HYPOTHESIS

Given the importance of the ability to predict inflation to monetary policy, this paper aims to develop a model that can forecast inflation and unemployment.

The hypotheses of this paper are the following:

- Inflation and Unemployment can be modeled independently through separate time series models
- There is a relationship between inflation and unemployment that is downward sloping → increase in unemployment leads to drop in inflation
- The forecast accuracy of the combined model is at par or better than the independent models

DATA

There are two time series datasets being used for this paper:

1. Monthly inflation data defined as: percentage change from a year ago of the "Consumer Price Index for all urban consumers: All items in US City average"[3], for the period Jan 1959 to Dec 2019
2. Monthly unemployment data for the period Jan 1959 to Dec 2019[4]

METHODOLOGY

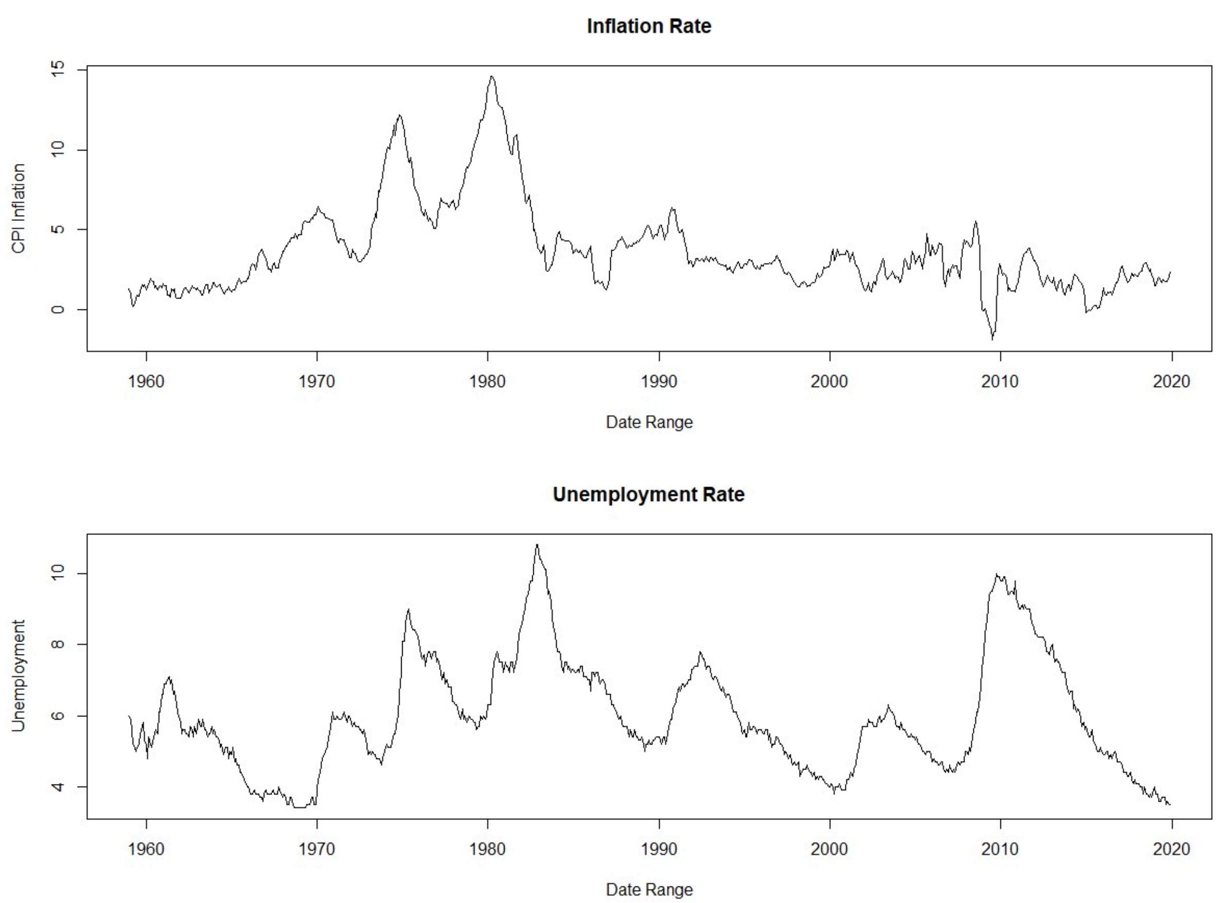
1. The first step is to plot the datasets separately for the full period and identify periods where trend changes
2. Split the data at these time periods so the trend can be captured and analyze these datasets separately

3. Divide the datasets into train and test datasets – the last year of the dataset (12 records) constitute the test dataset
4. Model the individual datasets based on appropriate SARIMA models and evaluate residuals to ensure they do not exhibit any autocorrelation or heteroskedasticity
5. Derive the 12-month ahead forecast and compare against the test datasets to evaluate model predictive power
6. Check whether the combined dataset exhibits cointegration and develop a VAR model and check whether there is a relationship
7. Derive the 12-month ahead forecast from the combined and compare against the individual models

ANALYSIS AND RESULTS

Plot the datasets separately for Inflation and Unemployment for the full time period. The plot is shown below in Figure 1.

Figure 1: Plot of Inflation Rate and Unemployment Rate



As can be seen from Figure 1, inflation has a different trend before and after 1980. Until 1980, the inflation rate is going up. After 1980, the inflation rate rapidly falls and remains stable for the rest of the time. Unemployment has a similar trend spiking at 1980 and then dropping in cycles until 2007. Post 2007 is the financial crisis which sees another sharp spike in unemployment and steady drop after that. This period is excluded for the rest of the analysis.

Inflation Dataset (1959-1980):

Figure 2 shows the plot of inflation from period Jan 1959 to Dec 1979 after decomposing into seasonality and trend components. Figure 3 shows the ACF and PACF of the inflation data. The very gradual decay of the ACF plot suggests that the data is not stationary.

Figure 2: Decomposition of Inflation time series from Jan 1959 to Dec 1979

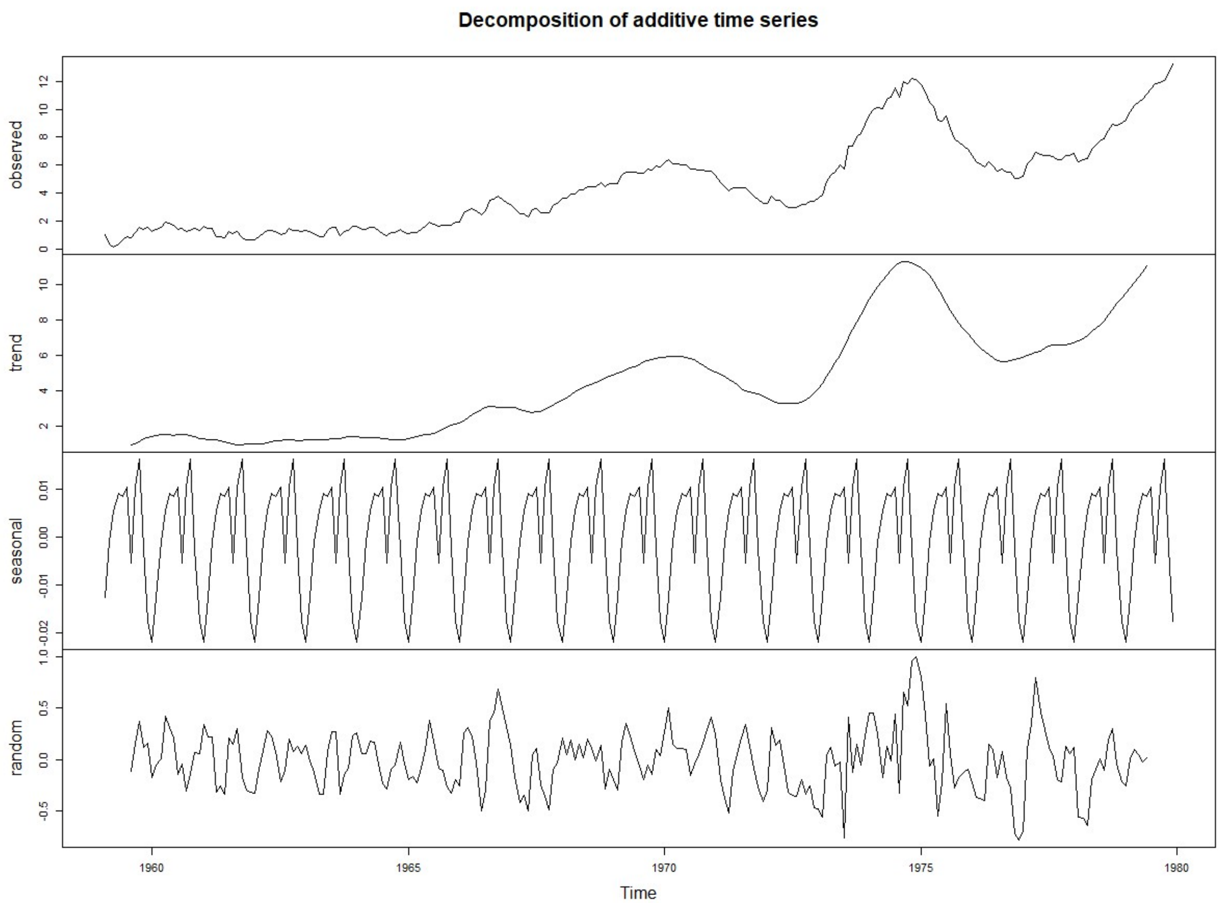
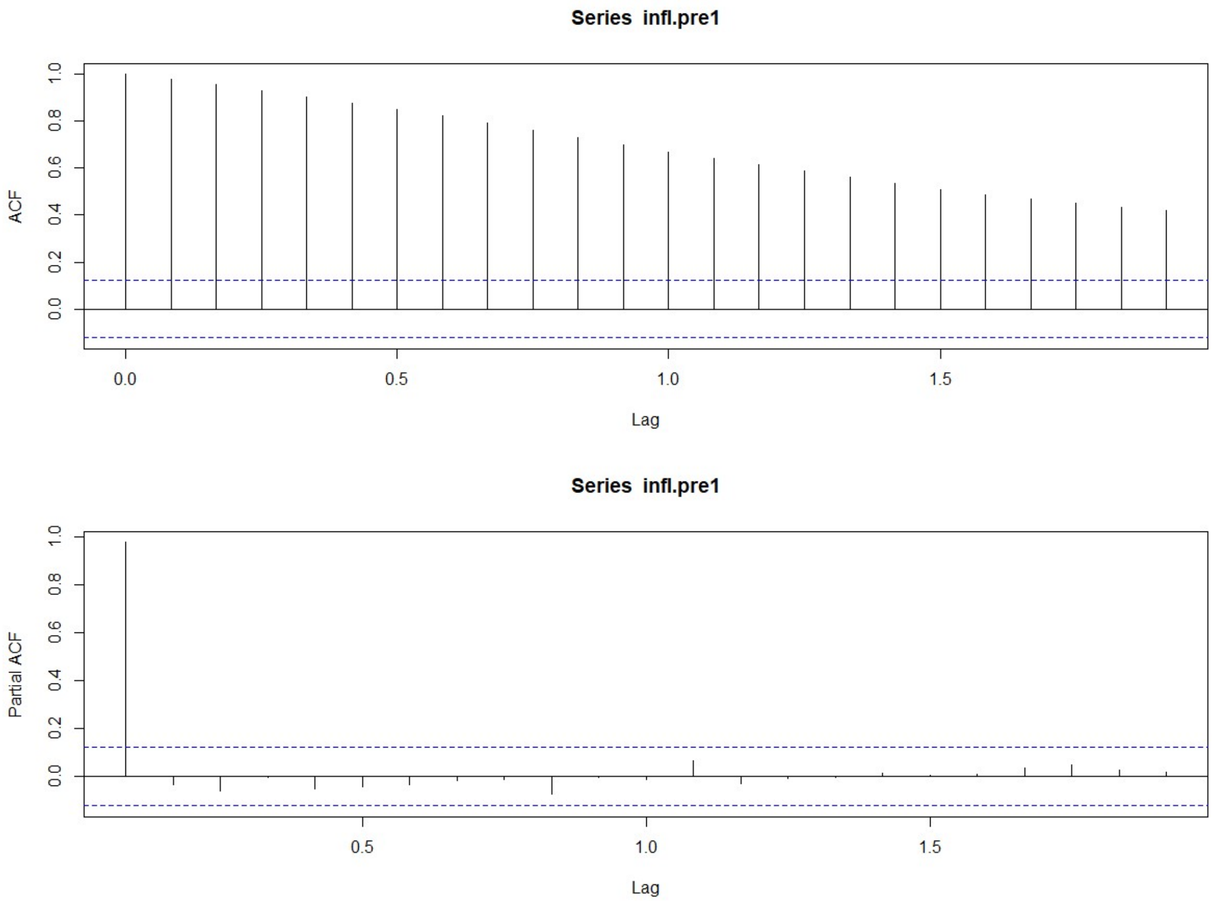


Figure 3: ACF/PACF plot of inflation time series from Jan 1959 to Dec 1979



We now develop a SARIMA model with a max order of 2 for all the AR, MA and differencing and seasonal components. The SARIMA model is shown below in Table 1.

Table 1: SARIMA model outcome for inflation time series from Jan 1959 to Dec 1979

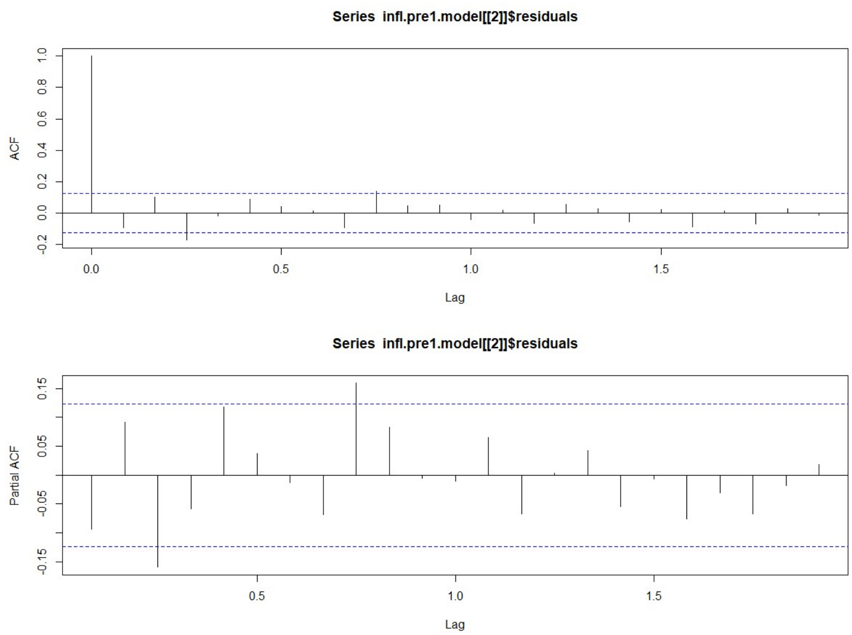
```
Call:
arima(x = x.ts, order = c(p, d, q), seasonal = list(order = c(P, D, Q), frequ
ency(x.ts)),
    method = "CSS")

Coefficients:
          ma1     sar1     sma1     sma2
      -0.7512  -0.3403  -0.5712  -0.3534
s.e.   0.0403   0.0953   0.1114   0.1079

sigma^2 estimated as 0.05322:  part log likelihood = 11.89
```

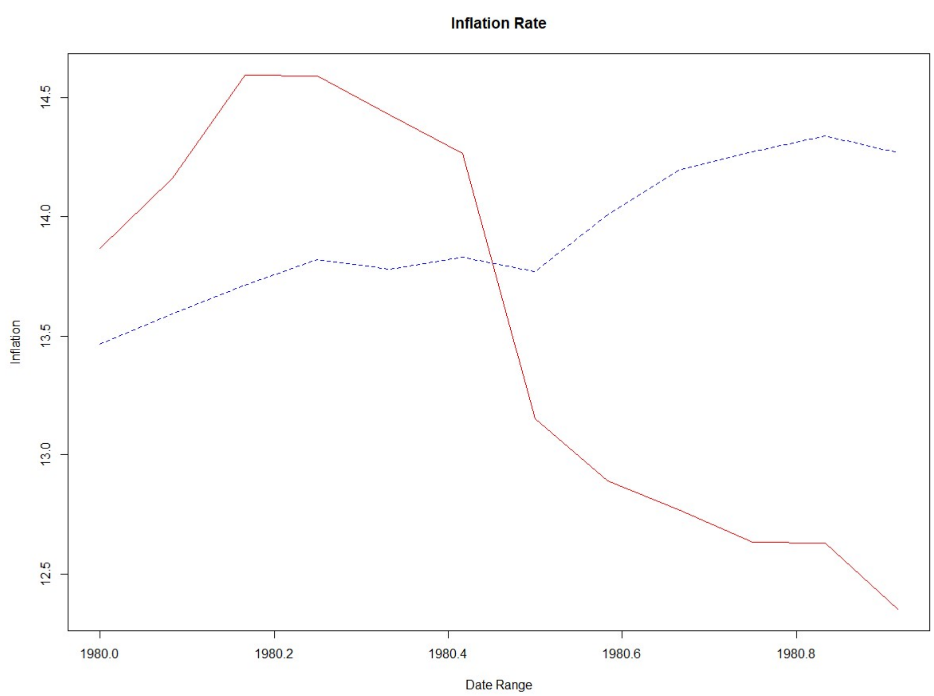
The model, as can be seen is a SARIMA (0,0,1,1,0,2). The ACF of the residuals are shown below in Figure 4.

Figure 4: ACF and PACF of residual from SARIMA model for Inflation from Jan 1959 – Dec 1979



As can be seen from the ACF/PACF plots, the residual exhibits no autocorrelation. We now predict the 12 period ahead forecast for the inflation in 1980 and compare against the actual inflation for the period. The plot is shown below in Figure 5. The accuracy is estimated by calculating the "Mean Absolute Percentage Error (MAPE)". The MAPE for this model is **7.2%**.

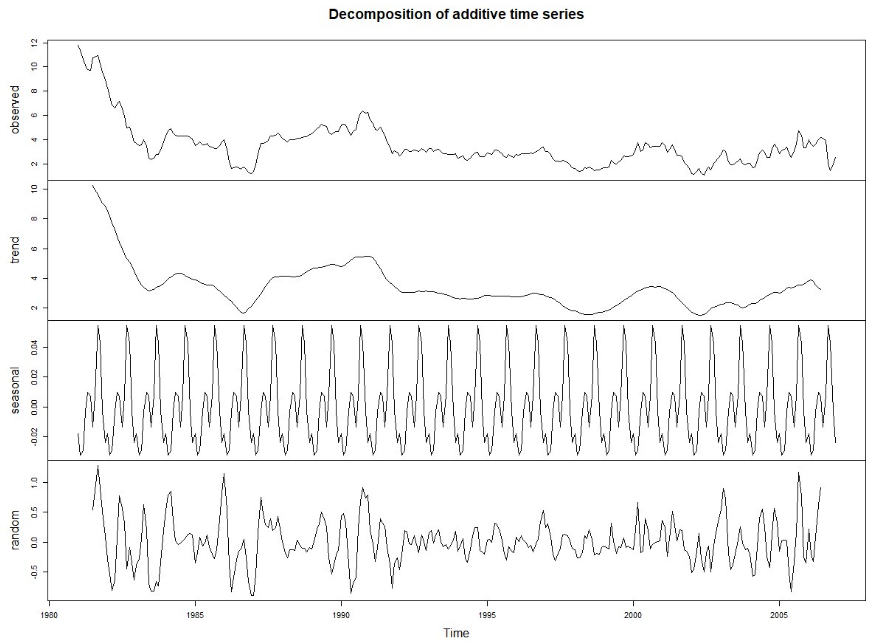
Figure 5: Actual (Red) vs Forecast (Blue) inflation for Jan 1980-Dec 1980



Inflation Dataset (1981-2007):

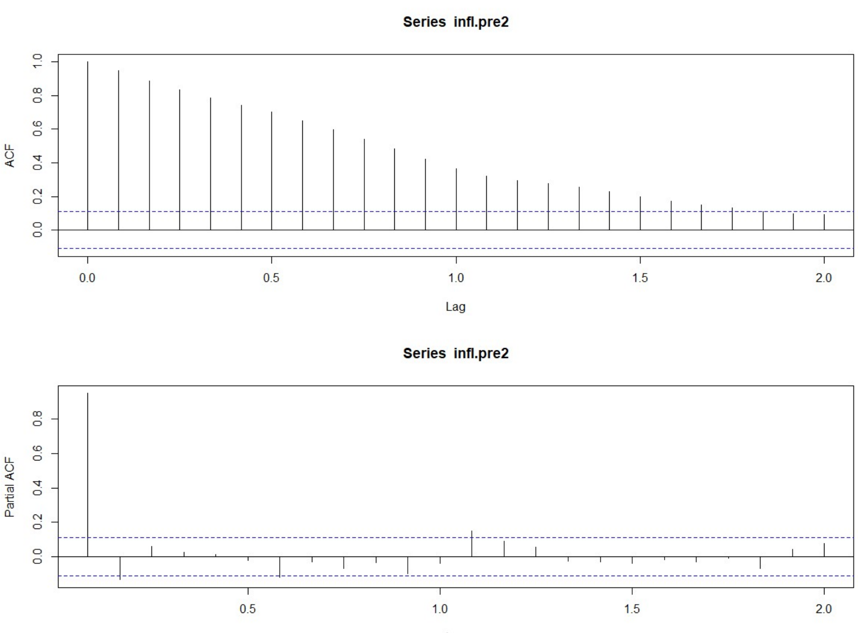
Figure 6 shows the plot of inflation from period Jan 1981 to Dec 2006 after decomposing into seasonal and trend components.

Figure 6: Decomposition of inflation time series from Jan 1981 – Dec 2006



ACF and PACF plots are shown below in Figure 7.

Figure 7: ACF and PACF plots for inflation from Jan 1981 – Dec 2006



From the ACF plot which decays very slowly, we can see that the series is not stationary. We develop a SARIMA model for this dataset and the model is shown below in Table 2.

Table 2: SARIMA model of inflation time series from Jan 1981 – Dec 2006

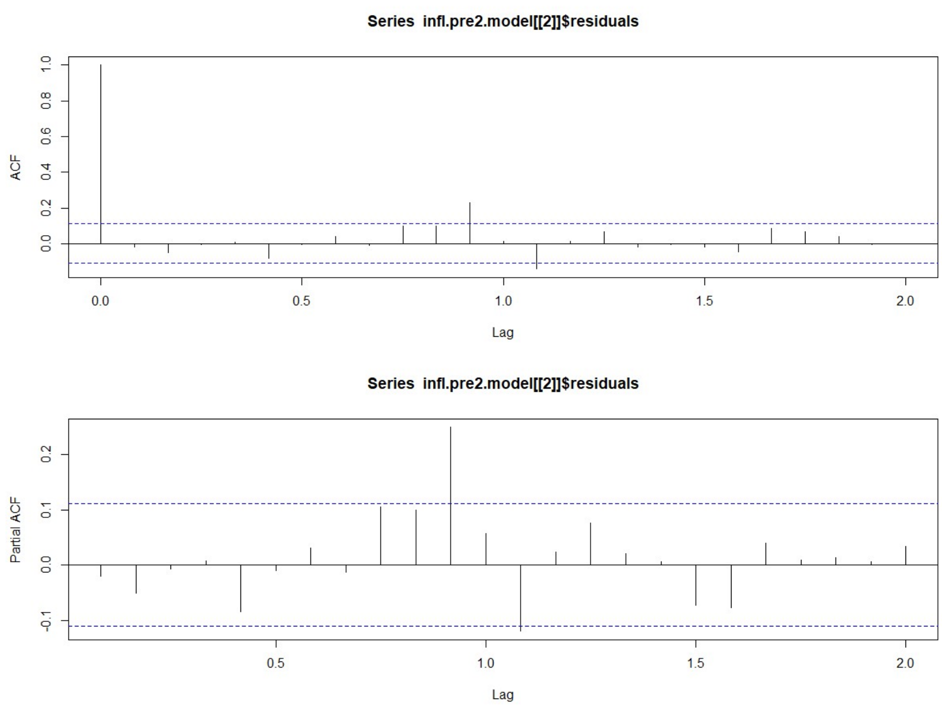
```
Call:
arima(x = x.ts, order = c(p, d, q), seasonal = list(order = c(P, D, Q), frequ
ency(x.ts)),
    method = "CSS")

Coefficients:
         ar1      ma1      ma2     sar1     sar2     sma1     sma2
      0.8738  -0.3665  -0.4152  -0.6985  -0.2791  -0.3413  -0.3099
s.e.  0.0261   0.0579   0.0509   0.1107   0.0589   0.1104   0.1046

sigma^2 estimated as 0.04697:  part log likelihood = 34.25
```

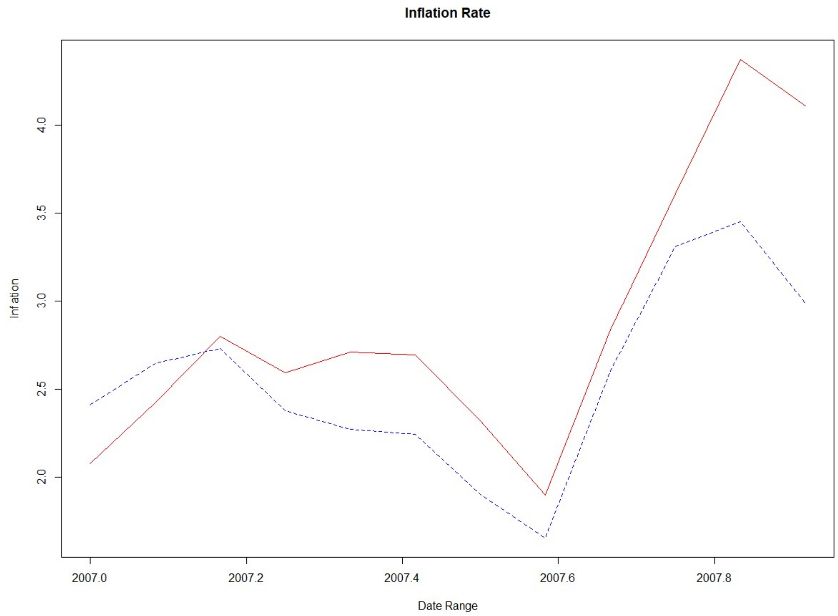
The model developed is a SARIMA (1,0,2,2,0,2). The ACF/PACF of the residuals for this SARIMA model is shown below in Figure 8. We can see that there is no evidence of autocorrelation in the residuals.

Figure 8: ACF/PACF of residuals for SARIMA model on inflation data for Jan 1981 – Dec 2006



The 12 month ahead forecast for Jan 2007 to Dec 2007 is shown below compared against the actual inflation in Figure 9. The MAPE for this model is **16%**.

Figure 9: Forecast vs Actual inflation for Jan 2007 – Dec 2007



Unemployment Dataset (1959-1980):

Figure 10 shows the plot of unemployment from period Jan 1959 to Dec 1979 after decomposing into seasonal and trend components.

Figure 10: Decomposition of unemployment from Jan 1959 – Dec 1979

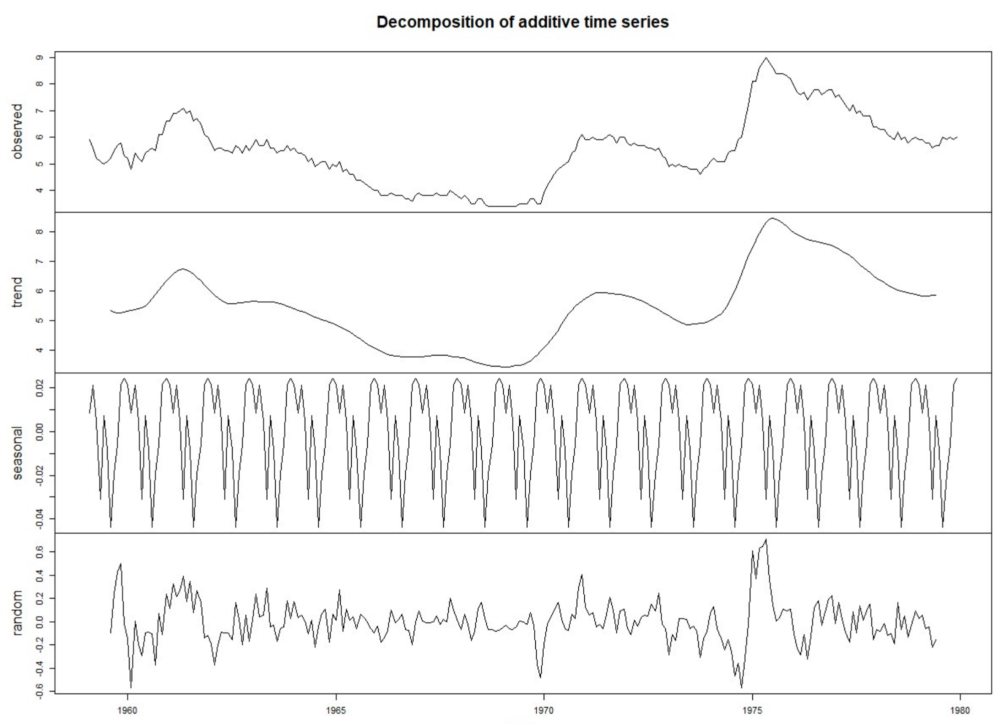
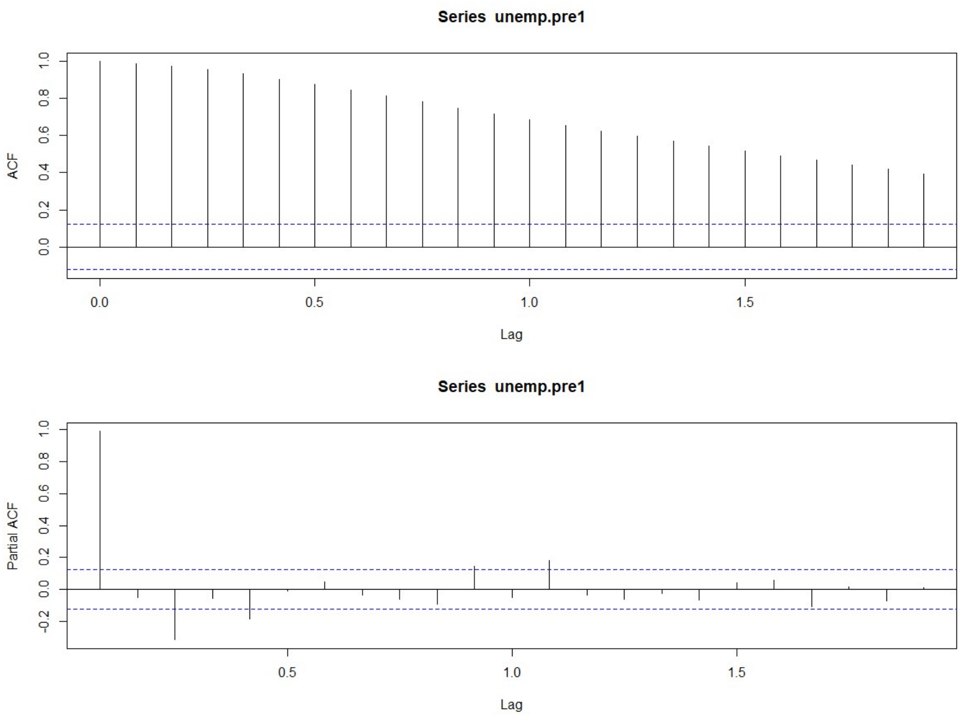


Figure 11: ACF/PACF of the unemployment time series from Jan 1959 – Dec 1979



From Figure 11 above, we can see that the series is not stationary. We develop a SARIMA model for this and the outcome is summarized in Table 3 below.

Table 3: SARIMA model for unemployment from Jan 1959 – Dec 1979

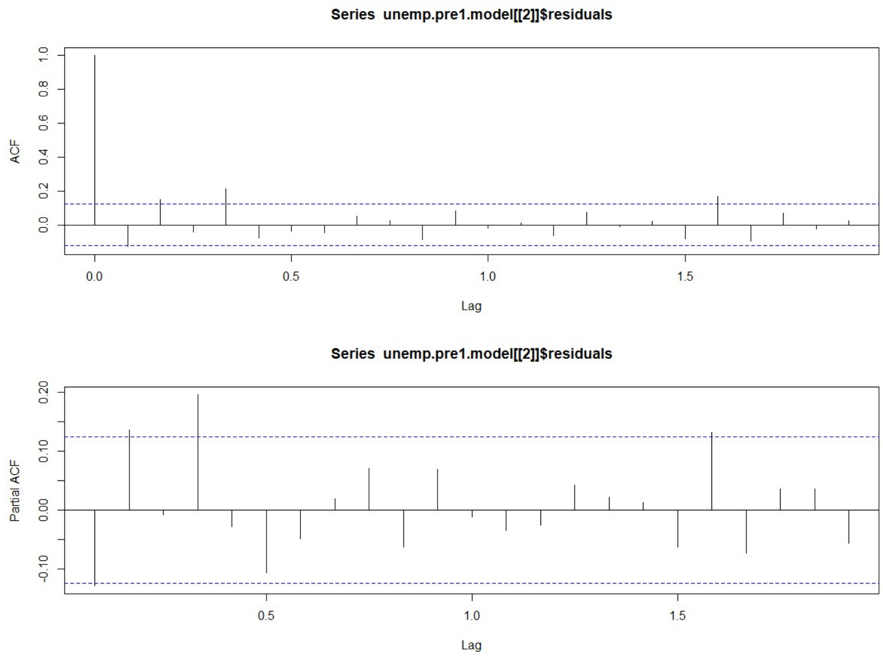
```
Call:
arima(x = x.ts, order = c(p, d, q), seasonal = list(order = c(P, D, Q), frequ
ency(x.ts)),
    method = "CSS")

Coefficients:
         ar1      ma1     sar1
      0.8351  -0.6763  -0.2289
s.e.  0.0691   0.0832   0.0610

sigma^2 estimated as 0.03249:  part log likelihood = 73.61
```

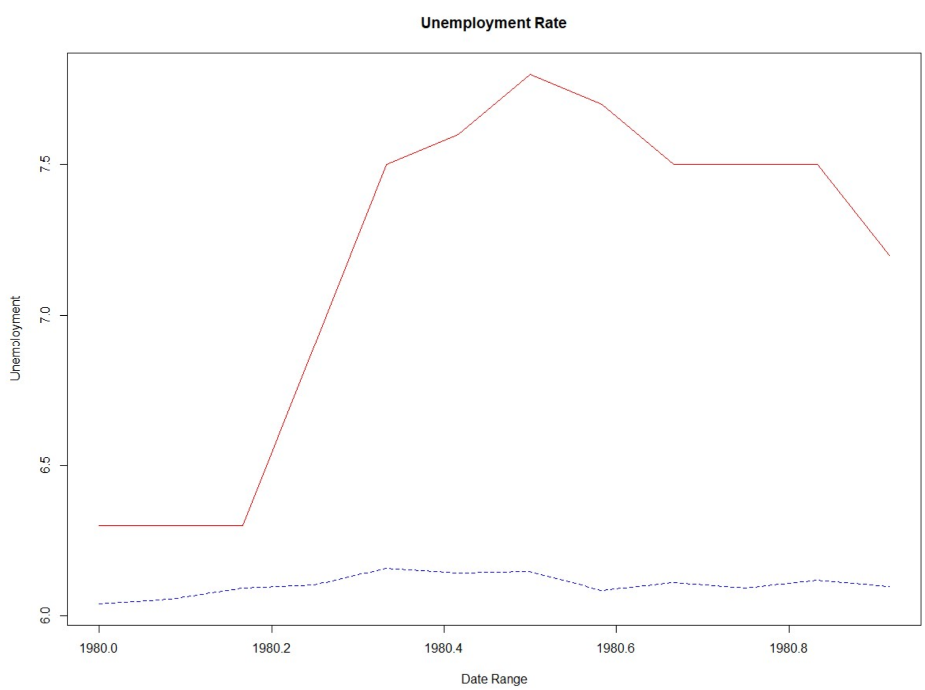
The model is SARIMA (1,0,1,1,0,0). The residuals of this model are checked for autocorrelation. From the output seen in Figure 12, we can conclude that autocorrelation does not exist.

Figure 12: ACF/PACF plot for SARIMA model residual for unemployment from Jan 1959 – Dec 1979



Forecast vs actual for unemployment for the 12 months of 1980 is shown below in Figure 13. The MAPE for this model is **17.5%**.

Figure 13: Forecast vs actual plot for Unemployment from Jan-Dec 1980



Unemployment Dataset (1981-2007):

Similarly, we develop a SARIMA model for Unemployment from Jan 1981 – Dec 2006 and forecast it for Jan 2007 – Dec 2007. The model is summarized below in Table 4 and the ACF/PACF for residuals from the model is shown in Figure 14 and the forecast vs actual in Figure 15. The MAPE for this model is **6.8%**.

Table 4: SARIMA model for unemployment from Jan 1981 – Dec 2006

```
Call:
arima(x = x.ts, order = c(p, d, q), seasonal = list(order = c(P, D, Q), frequ
ency(x.ts)),
    method = "CSS")

Coefficients:
         ar1      ma1     sar1
      0.8909  -0.7901  -0.0563
s.e.  0.0364   0.0416   0.0556

sigma^2 estimated as 0.0218:  part log likelihood = 153.62
```

The model is SARIMA (1,0,1,1,0,0).

Figure 14: ACF/PACF of the residuals from SARIMA model for unemployment from Jan 1981 – Dec 2006

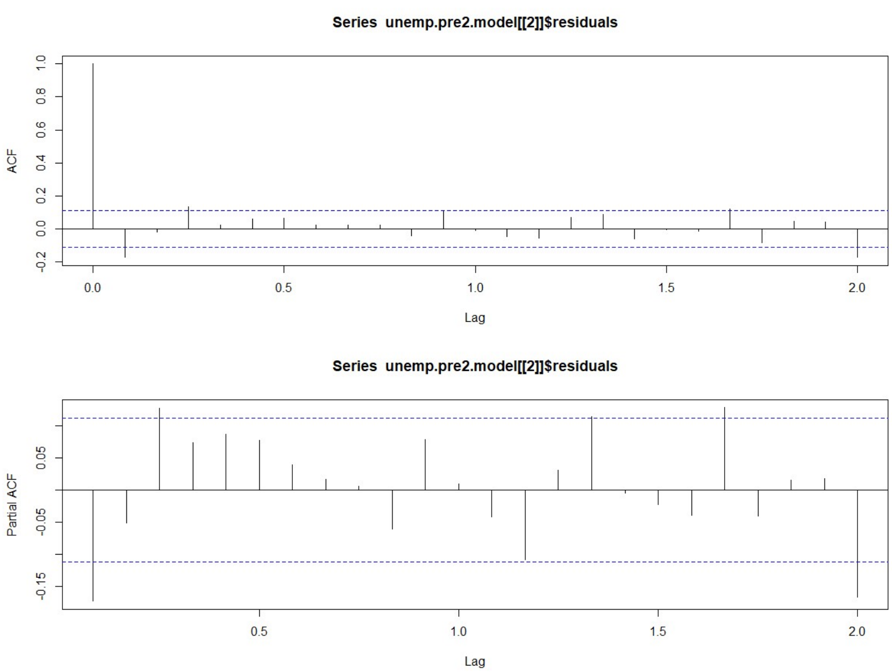
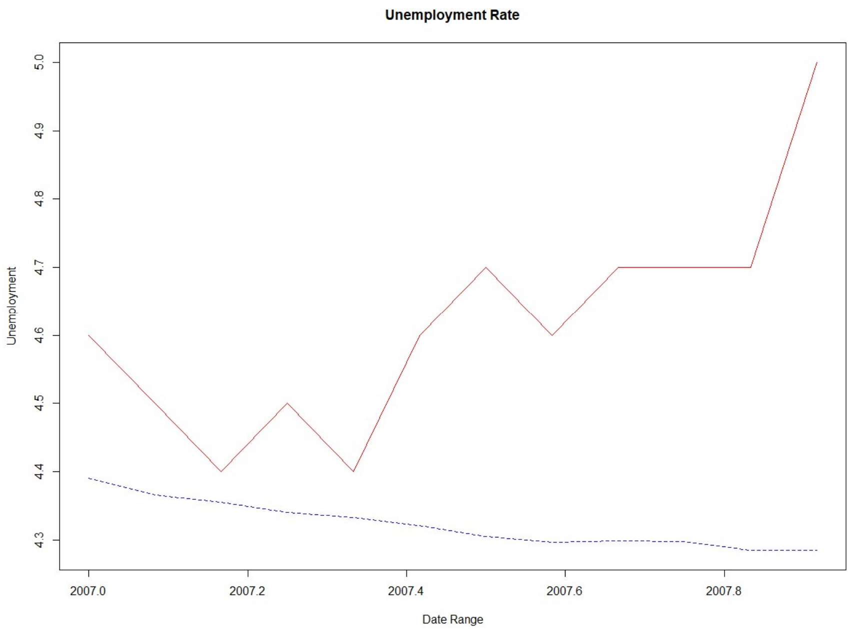


Figure 15: Forecast vs actual for unemployment from Jan – Dec 2007



<u>Combined Unemployment vs Inflation dataset (1959-1980)</u>:

We first examine whether unit roots exist in the inflation and unemployment datasets in Table 5 below.

Table 5: ADF test for inflation and unemployment datasets from 1959-1980

Augmented Dickey-Fuller Test For **Unemployment** alternative: stationary Type 1: no drift no trend lag ADF p.value [1,] 0 -0.240 0.575 [2,] 1 -0.157 0.599 [3,] 2 -0.146 0.602 [4,] 3 -0.166 0.596 [5,] 4 -0.246 0.573 Type 2: with drift no trend lag ADF p.value [1,] 0 -1.18 0.634 [2,] 1 -1.23 0.616 [3,] 2 -1.75 0.425 [4,] 3 -1.85 0.388 [5,] 4 -2.21 0.246 Type 3: with drift and trend lag ADF p.value [1,] 0 -1.46 0.804 [2,] 1 -1.44 0.811 [3,] 2 -1.85 0.638 [4,] 3 -1.95 0.597 [5,] 4 -2.32 0.443 ----	Augmented Dickey-Fuller Test For **Inflation** alternative: stationary Type 1: no drift no trend lag ADF p.value [1,] 0 2.550 0.990 [2,] 1 1.960 0.988 [3,] 2 1.218 0.942 [4,] 3 1.319 0.953 [5,] 4 0.975 0.911 Type 2: with drift no trend lag ADF p.value [1,] 0 1.1006 0.990 [2,] 1 0.4805 0.984 [3,] 2 -0.1918 0.933 [4,] 3 0.0425 0.959 [5,] 4 -0.2087 0.930 Type 3: with drift and trend lag ADF p.value [1,] 0 -0.758 0.964 [2,] 1 -1.102 0.921 [3,] 2 -1.845 0.641 [4,] 3 -1.735 0.687 [5,] 4 -2.210 0.487 ----

We can see that all the p-values for both series are high enough that there is a unit root for both series. So we take first difference and conduct ADF test again and find that they are stationary as can be seen in Table 6.

Table 6: ADF test of first differenced unemployment and inflation series

```
Augmented Dickey-Fuller Test
For Unemployment (first diff)
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 2.550   0.990
[2,]   1 1.960   0.988
[3,]   2 1.218   0.942
[4,]   3 1.319   0.953
[5,]   4 0.975   0.911
Type 2: with drift no trend
     lag     ADF p.value
[1,]   0  1.1006   0.990
[2,]   1  0.4805   0.984
[3,]   2 -0.1918   0.933
[4,]   3  0.0425   0.959
[5,]   4 -0.2087   0.930
Type 3: with drift and trend
     lag    ADF p.value
[1,]   0 -0.758   0.964
[2,]   1 -1.102   0.921
[3,]   2 -1.845   0.641
[4,]   3 -1.735   0.687
[5,]   4 -2.210   0.487
----
```

```
Augmented Dickey-Fuller Test
For Inflation (first diff)
alternative: stationary

Type 1: no drift no trend
     lag    ADF p.value
[1,]   0 -14.84    0.01
[2,]   1  -7.84    0.01
[3,]   2  -6.65    0.01
[4,]   3  -5.17    0.01
[5,]   4  -4.82    0.01
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -14.81    0.01
[2,]   1  -7.82    0.01
[3,]   2  -6.64    0.01
[4,]   3  -5.17    0.01
[5,]   4  -4.82    0.01
Type 3: with drift and trend
     lag    ADF p.value
[1,]   0 -14.78    0.01
[2,]   1  -7.80    0.01
[3,]   2  -6.63    0.01
[4,]   3  -5.15    0.01
[5,]   4  -4.80    0.01
----
```

We now conduct Phillips-Ouliaris test to check whether the two series are cointegrated. The results are shown below in Table 7.

Table 7: Cointegration test for inflation and unemployment

```
        Phillips-Ouliaris Cointegration Test

data:  cbind(infl.pre1, unemp.pre1)
Phillips-Ouliaris demeaned = 0.25086, Truncation lag parameter = 2, p-value =
0.15
```

From the cointegration test p-value we can see that we cannot reject the null hypothesis that there is no cointegration in the dataset. We now build a VAR model with inflation and unemployment first differences to check whether there is a relationship between the two. The VAR is shown below in Table 8.

Table 8: VAR of inflation and unemployment with lag 3, trend and constant

```
$dinfl.pre1
                  Estimate   Std. Error     t value     Pr(>|t|)
dinfl.pre1.l1   0.1381474036 0.0644173537  2.14456813 3.299622e-02
dunemp.pre1.l1  0.0048725595 0.1063109824  0.04583308 9.634816e-01
dinfl.pre1.l2   0.3139150192 0.0616022919  5.09583344 7.054191e-07
dunemp.pre1.l2 -0.2020575836 0.0992473292 -2.03589946 4.286345e-02
dinfl.pre1.l3  -0.0686884911 0.0635515320 -1.08083140 2.808615e-01
dunemp.pre1.l3 -0.1498451318 0.1058653286 -1.41543160 1.582437e-01
const          -0.0083562748 0.0395708185 -0.21117265 8.329325e-01
trend           0.0003313341 0.0002736491  1.21079916 2.271684e-01

$dunemp.pre1
                  Estimate   Std. Error    t value     Pr(>|t|)
dinfl.pre1.l1  -3.545943e-02 0.0389991520 -0.9092359 3.641416e-01
dunemp.pre1.l1  9.718815e-03 0.0643621312  0.1510021 8.801015e-01
dinfl.pre1.l2   4.140978e-02 0.0372948749  1.1103344 2.679708e-01
dunemp.pre1.l2  3.310123e-01 0.0600856983  5.5090028 9.311509e-08
dinfl.pre1.l3   5.402589e-02 0.0384749717  1.4041827 1.615626e-01
dunemp.pre1.l3  4.623065e-02 0.0640923262  0.7213134 4.714218e-01
const           5.074575e-03 0.0239567178  0.2118226 8.324259e-01
trend          -3.576389e-05 0.0001656709 -0.2158731 8.292708e-01
```

As can be seen from Table 8, the equation for inflation shows that the coefficients of unemployment are not statistically significant and similarly, the equation for unemployment shows that the coefficients for inflation are not statistically significant. Therefore, these two time series are not cointegrated or dependent on each other in a statistically significant way for this time period.

Combined Unemployment vs Inflation dataset (1981-2007):

We first examine whether unit roots exist in the inflation and unemployment datasets in Table 9 below.

Table 9: ADF test for unemployment and inflation from Jan 1981 – Dec 2006

```
Augmented Dickey-Fuller Test
For Unemployment
alternative: stationary

Type 1: no drift no trend
     lag    ADF p.value
[1,]   0 -1.277   0.223
[2,]   1 -1.191   0.253
[3,]   2 -1.082   0.292
[4,]   3 -0.888   0.361
[5,]   4 -1.006   0.319
[6,]   5 -1.021   0.314
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -0.817   0.762
[2,]   1 -0.841   0.754
[3,]   2 -1.039   0.684
[4,]   3 -1.366   0.569
[5,]   4 -1.705   0.444
[6,]   5 -1.975   0.338
```

```
Augmented Dickey-Fuller Test
For Inflation
alternative: stationary

Type 1: no drift no trend
     lag   ADF p.value
[1,]   0 -3.52    0.01
[2,]   1 -2.83    0.01
[3,]   2 -2.84    0.01
[4,]   3 -2.73    0.01
[5,]   4 -2.61    0.01
[6,]   5 -2.61    0.01
Type 2: with drift no trend
     lag   ADF p.value
[1,]   0 -4.79    0.01
[2,]   1 -4.49    0.01
[3,]   2 -4.14    0.01
[4,]   3 -4.07    0.01
[5,]   4 -3.95    0.01
[6,]   5 -4.02    0.01
```

```
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -1.42   0.823
[2,]   1 -1.51   0.785
[3,]   2 -1.76   0.679
[4,]   3 -2.35   0.429
[5,]   4 -2.64   0.307
[6,]   5 -3.04   0.139
----
```

```
Type 3: with drift and trend
     lag   ADF p.value
[1,]   0 -4.35  0.0100
[2,]   1 -4.45  0.0100
[3,]   2 -3.93  0.0124
[4,]   3 -3.91  0.0137
[5,]   4 -3.80  0.0191
[6,]   5 -3.88  0.0148
----
```

From Table 9, we can see that the unemployment series has unit root, but the inflation series does not.

We look at ADF test for first differenced series as shown in Table 10. We can see that both series do not have a unit root after first differencing.

Table 10: ADF test for first difference of unemployment and inflation from Jan 1981 – Dec 2006

```
Augmented Dickey-Fuller Test
For unemployment (first difference)
alternative: stationary

Type 1: no drift no trend
     lag    ADF p.value
[1,]   0 -16.63    0.01
[2,]   1 -10.21    0.01
[3,]   2  -6.71    0.01
[4,]   3  -5.38    0.01
[5,]   4  -4.48    0.01
[6,]   5  -3.99    0.01
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -16.66    0.01
[2,]   1 -10.24    0.01
[3,]   2  -6.73    0.01
[4,]   3  -5.41    0.01
[5,]   4  -4.51    0.01
[6,]   5  -4.01    0.01
Type 3: with drift and trend
     lag    ADF p.value
[1,]   0 -16.63    0.01
[2,]   1 -10.23    0.01
[3,]   2  -6.72    0.01
[4,]   3  -5.40    0.01
[5,]   4  -4.50    0.01
[6,]   5  -4.01    0.01
----
```

```
Augmented Dickey-Fuller Test
For inflation (first difference)
alternative: stationary

Type 1: no drift no trend
     lag    ADF p.value
[1,]   0 -11.96    0.01
[2,]   1 -12.55    0.01
[3,]   2  -9.96    0.01
[4,]   3  -8.63    0.01
[5,]   4  -7.68    0.01
[6,]   5  -6.32    0.01
Type 2: with drift no trend
     lag    ADF p.value
[1,]   0 -11.99    0.01
[2,]   1 -12.60    0.01
[3,]   2 -10.01    0.01
[4,]   3  -8.69    0.01
[5,]   4  -7.74    0.01
[6,]   5  -6.41    0.01
Type 3: with drift and trend
     lag    ADF p.value
[1,]   0 -12.08    0.01
[2,]   1 -12.71    0.01
[3,]   2 -10.13    0.01
[4,]   3  -8.80    0.01
[5,]   4  -7.86    0.01
[6,]   5  -6.60    0.01
----
```

We now check for cointegration between the two first differenced series. The results of the cointegration test are shown below in Table 11.

Table 11: Cointegration test between inflation and unemployment for Jan 1981 – Dec 2006

```
        Phillips-Ouliaris Cointegration Test

data:  cbind(dinfl.pre2, dunemp.pre2)
Phillips-Ouliaris demeaned = -183.29, Truncation lag parameter = 3, p-value =
0.01
```

From the p-value, we can conclude that the two series are cointegrated. Therefore, we can build a VAR model with the data series. The VAR model is shown below in Table 12.

Table 12: VAR model for inflation vs unemployment from Jan 1981 – Dec 2006

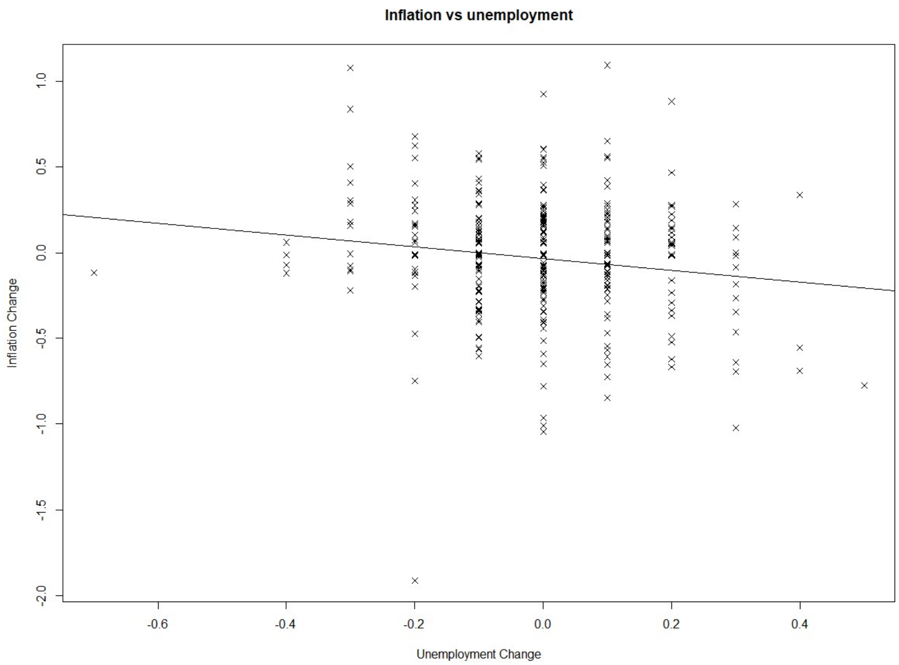
```
$infl.pre2
                  Estimate   Std. Error    t value     Pr(>|t|)
infl.pre2.l1   1.3486369247 0.0557112824 24.2076087 1.296327e-72
unemp.pre2.l1 -0.1657893360 0.1171297877 -1.4154327 1.579755e-01
infl.pre2.l2  -0.6403211234 0.0881509928 -7.2639128 3.239165e-12
unemp.pre2.l2 -0.1360005960 0.1623910061 -0.8374885 4.029825e-01
infl.pre2.l3   0.2512669739 0.0543833561  4.6202918 5.689029e-06
unemp.pre2.l3  0.2612356390 0.1167618171  2.2373379 2.599599e-02
const          0.4549413062 0.1677694293  2.7117056 7.078541e-03
trend         -0.0005920881 0.0003302163 -1.7930311 7.397185e-02

$unemp.pre2
                  Estimate   Std. Error    t value     Pr(>|t|)
infl.pre2.l1   0.0231976296 0.0271440112  0.8546132 3.934450e-01
unemp.pre2.l1  0.9582639849 0.0570687324 16.7914012 4.122158e-45
infl.pre2.l2  -0.0410099188 0.0429494967 -0.9548405 3.404246e-01
unemp.pre2.l2  0.1196586778 0.0791211959  1.5123467 1.314949e-01
infl.pre2.l3   0.0486869086 0.0264970104  1.8374491 6.712923e-02
unemp.pre2.l3 -0.0914120548 0.0568894474 -1.6068368 1.091387e-01
const         -0.0610128275 0.0817417060 -0.7464100 4.560023e-01
trend          0.0001604723 0.0001608901  0.9974032 3.193700e-01
```

From the equation for inflation, we can see that the coefficient for Unemployment with lag 3 is statistically significant at 95% confidence level. Similarly, we can see that inflation of lag 3 is statistically significant at 90% confidence level in the unemployment equation.

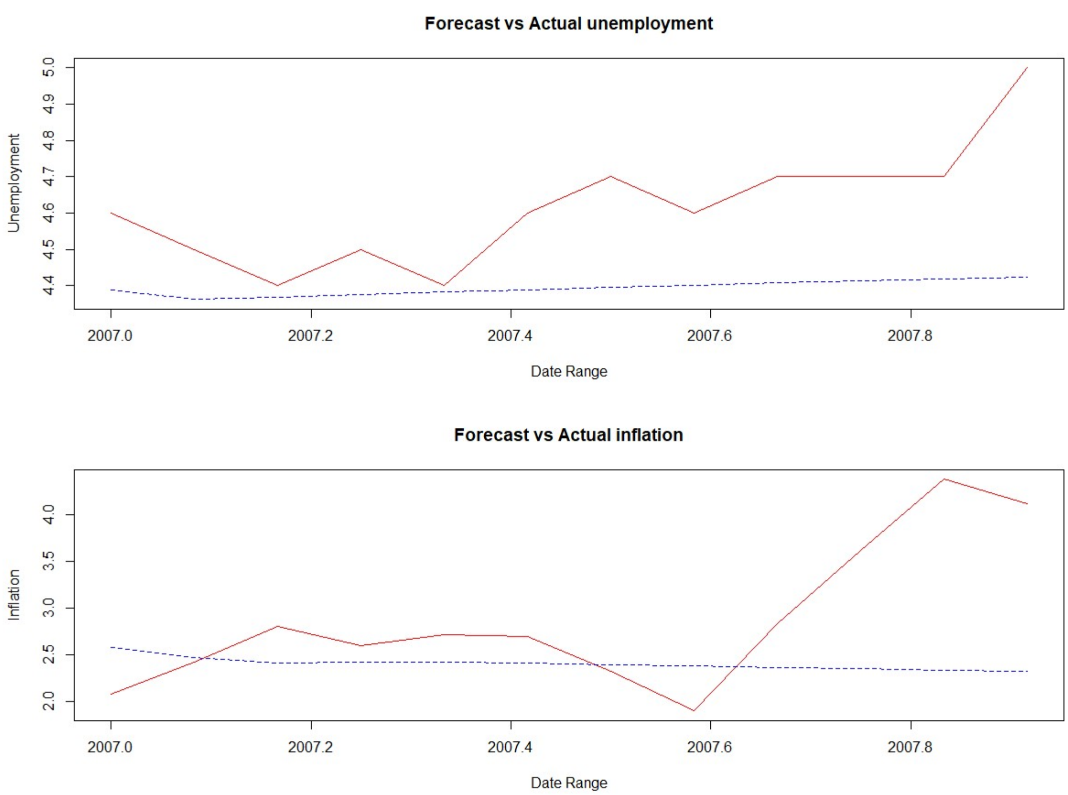
We can also see there is some evidence of correlation between the two time series as seen in the Figure 16. The correlation between the two variables is -0.15.

Figure 16: Relationship between change in unemployment and change in inflation



We can see the forecast vs actual inflation and unemployment from the VAR model in Figure 17 below.

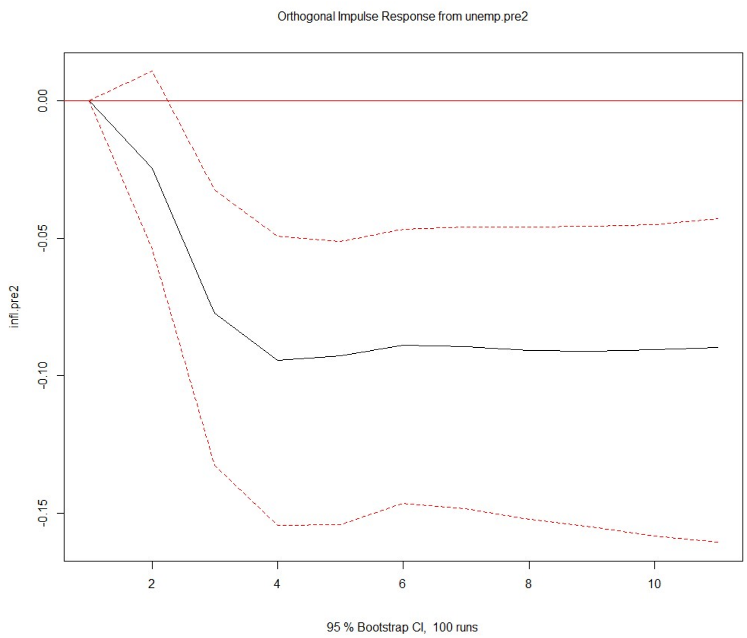
Figure 17: Forecast vs Actual Unemployment and Inflation



The MAPE for the unemployment fit is **5.1%**, the MAPE for inflation is **27.4%**.

The relationship between inflation and unemployment from the VAR model can be understood through the impulse response function of inflation to an impulse of unemployment. This can be seen in Figure 18.

Figure 18: IRF of inflation to an impulse in unemployment



From the Figure, we can see that for a 1% increase in unemployment, inflation drops by the highest to 9.5% at lag 4.

CONCLUSIONS

From the analysis conducted in this paper, we can conclude the following:

- Inflation can be modeled as:
    - SARIMA (0,0,1,1,0,2) from Jan 1959 – Dec 1980
    - SARIMA (1,0,2,2,0,2) from Jan 1981 – Dec 2007
- Unemployment can be modeled as:
    - SARIMA (1,0,1,1,0,0) from Jan 1959 – Dec 1980
    - SARIMA (1,0,1,1,0,0) from Jan 1981 – Dec 2007
- There is a relationship between inflation and unemployment that is downward sloping for the period Jan 1981 – Dec 2007 → increase in unemployment leads to drop in inflation
    - A 1% increase in unemployment results in a 9.5% drop in inflation at lag 4
- The forecast accuracy of the combined model vs standalone model as defined by MAPE:
    - MAPE for inflation – Combined: 27.4%, Individual: 16%
    - MAPE for unemployment – Combined: 5.1%, Individual: 6.8%

REFERENCES

[1] Phillips, A. (1958). The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957. Economica, 25(100), new series, 283-299. doi:10.2307/2550759
[2] Brian Reinbold and Yi Wen, "Is the Phillips Curve Still Alive?," Federal Reserve Bank of St. Louis *Review*, Second Quarter 2020, pp. 121-44. https://doi.org/10.20955/r.102.121-44

[3] U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCSL, August 1, 2020.
[4] U.S. Bureau of Labor Statistics, Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/UNRATE, August 1, 2020.

APPENDIX 1: PROJECT CODE

```
# Project Code

# Load all the libraries we need

library(MLmetrics)

library(vars)

library(aTSA)


# Set working directory

setwd("C:/Users/subha/Desktop/ANLY565/RScript")


# Load Inflation and unemployment data

infl_unemp<-read_excel("INFL_UNRATE.xlsx", col_types = c("date", "numeric", "numeric"))



date<-infl_unemp$DATE

infl<-infl_unemp$INFL

unemp <- infl_unemp$UNRATE


inflts<-ts(infl, start=c(1959,1), freq=12)

unempts<-ts(unemp, start=c(1959,1), freq=12)
```

```
# Plot the full datasets
par(mfrow=c(2,1))
plot(inflts, xlab="Date Range", ylab="CPI Inflation", main="Inflation Rate")
plot(unempts, xlab="Date Range", ylab="Unemployment", main="Unemployment Rate")

# Split the data into pre and post 1980 period
infl.pre1  <- window(inflts, start = c(1959,2), end = c(1979,12), freq = 12)
infl.post1 <- window(inflts, start = c(1980,1), end = c(1980,12), freq = 12)
infl.pre2  <- window(inflts, start = c(1981,1), end = c(2006,12), freq = 12)
infl.post2 <- window(inflts, start = c(2007,1), end = c(2007,12), freq = 12)



# Function to develop a SARIMA model with a max order of 2
get.best.sarima <- function(x.ts, maxord = c(1,1,1,1,1,1))
{
 best.aic <- 1e8
 n <- length(x.ts)
 for (p in 0:maxord[1]) for(d in 0:maxord[2]) for(q in 0:maxord[3])
  for (P in 0:maxord[4]) for(D in 0:maxord[5]) for(Q in 0:maxord[6])
  {
   fit <- arima(x.ts, order = c(p,d,q),
                seas = list(order = c(P,D,Q),
                       frequency(x.ts)), method = "CSS")
   fit.aic <- -2 * fit$loglik + (log(n) + 1) * length(fit$coef)
   if (fit.aic < best.aic)
   {
    best.aic <- fit.aic
    best.fit <- fit
```

```
        best.fit.coef <- fit$coef
        best.model <- c(p,d,q,P,D,Q)
      }
    }
  list(best.aic, best.fit, best.fit.coef, best.model)
}
```

```
infl.pre1.dec<-decompose(infl.pre1)
plot(infl.pre1.dec)

acf(infl.pre1)
pacf(infl.pre1)

infl.pre1.model<-get.best.sarima(infl.pre1, maxord = c(2,2,2,2,2,2))
infl.pre1.model[[2]]
# It is SARIMA(0,0,1,1,0,2)

acf(infl.pre1.model[[2]]$residuals)
pacf(infl.pre1.model[[2]]$residuals)

infl1.pred <- predict(infl.pre1.model[[2]], n.ahead=12)

par(mfrow=c(1,1))

ts.plot(infl.post1, infl1.pred$pred, lty = 1:2, col=c("red", "blue"), xlab="Date Range", ylab="Inflation",
main="Inflation Rate")
```

```
MAPE(infl.post1, infl1.pred$pred)

#Pre 2
infl.pre2.dec<-decompose(infl.pre2)
plot(infl.pre2.dec)

par(mfrow=c(2,1))
acf(infl.pre2)
pacf(infl.pre2)

infl.pre2.model<-get.best.sarima(infl.pre2, maxord = c(2,2,2,2,2,2))
infl.pre2.model[[2]]
# It is SARIMA(1,0,2,2,0,2)

acf(infl.pre2.model[[2]]$residuals)
pacf(infl.pre2.model[[2]]$residuals)

infl2.pred <- predict(infl.pre2.model[[2]], n.ahead=12)
par(mfrow=c(1,1))

ts.plot(infl.post2, infl2.pred$pred, lty = 1:2, col=c("red", "blue"), xlab="Date Range", ylab="Inflation",
main="Inflation Rate")

MAPE(infl.post2, infl2.pred$pred)

unemp.pre1  <- window(unempts, start = c(1959,2), end = c(1979,12), freq = 12)
```

```r
unemp.post1 <- window(unempts, start = c(1980,1), end = c(1980,12), freq = 12)
unemp.pre2  <- window(unempts, start = c(1981,1), end = c(2006,12), freq = 12)
unemp.post2 <- window(unempts, start = c(2007,1), end = c(2007,12), freq = 12)

unemp.pre1.dec<-decompose(unemp.pre1)
plot(unemp.pre1.dec)

par(mfrow=c(2,1))
acf(unemp.pre1)
pacf(unemp.pre1)

unemp.pre1.model<-get.best.sarima(unemp.pre1)
unemp.pre1.model[[2]]
# It is SARIMA(1,0,1,1,0,0)

acf(unemp.pre1.model[[2]]$residuals)
pacf(unemp.pre1.model[[2]]$residuals)

par(mfrow=c(1,1))
unemp1.pred<-predict(unemp.pre1.model[[2]], n.ahead=12)

ts.plot(unemp.post1, unemp1.pred$pred, lty=1:2, col=c("red", "blue"), xlab="Date Range",
ylab="Unemployment", main="Unemployment Rate")
MAPE(unemp.post1, unemp1.pred$pred)

par(mfrow=c(2,1))
acf(unemp.pre2)
pacf(unemp.pre2)
```

```
unemp.pre2.model<-get.best.sarima(unemp.pre2)

unemp.pre2.model[[2]]

# It is SARIMA (1,0,1,1,0,0)


par(mfrow=c(2,1))

acf(unemp.pre2.model[[2]]$residuals)

pacf(unemp.pre2.model[[2]]$residuals)


par(mfrow=c(1,1))

unemp2.pred<-predict(unemp.pre2.model[[2]], n.ahead=12)

ts.plot(unemp.post2, unemp2.pred$pred, lty=1:2, col=c("red", "blue"), xlab="Date Range",
ylab="Unemployment", main="Unemployment Rate")

MAPE(unemp.post2, unemp2.pred$pred)


adf.test(unemp.pre1)

adf.test(infl.pre1)


dunemp.pre1<-diff(unemp.pre1)

dinfl.pre1<-diff(infl.pre1)


adf.test(dunemp.pre1)

adf.test(dinfl.pre1)


po.test(cbind(infl.pre1, unemp.pre1))


inf_unemp.var <- VAR(cbind(dinfl.pre1, dunemp.pre1), p = 3, type = "both")

coef(inf_unemp.var)
```

```
adf.test(unemp.pre2)
adf.test(infl.pre2)

dunemp.pre2<-diff(unemp.pre2)
dinfl.pre2<-diff(infl.pre2)

adf.test(dunemp.pre2)
adf.test(dinfl.pre2)

po.test(cbind(dinfl.pre2, dunemp.pre2))

inf_unemp.var2 <- VAR(cbind(infl.pre2, unemp.pre2), p = 3, type = "both")
coef(inf_unemp.var2)

plot(dunemp.pre2, dinfl.pre2, pch = 4, main = "Inflation vs unemployment", xlab="Unemployment
Change", ylab="Inflation Change")
abline(lm(dinfl.pre2~dunemp.pre2))
cor(dunemp.pre2, dinfl.pre2)

infl_unemp2.pred<-predict(inf_unemp.var2, n.ahead=12)
par(mfrow=c(2,1))
```

```
ts.plot(unemp.post2, infl_unemp2.pred$fcst$unemp.pre2[,1], lty=1:2, col=c("red", "blue"), main =
"Forecast vs Actual unemployment", xlab="Date Range", ylab="Unemployment")

ts.plot(infl.post2, infl_unemp2.pred$fcst$infl.pre2[,1], lty=1:2, col=c("red", "blue"), main = "Forecast vs
Actual inflation", xlab="Date Range", ylab="Inflation")

MAPE(unemp.post2, infl_unemp2.pred$fcst$unemp.pre2[,1])

MAPE(infl.post2, infl_unemp2.pred$fcst$infl.pre2[,1])



irf(inf_unemp.var2, impulse = "unemp.pre2", response = c("infl.pre2"), boot = F)


plot(irf(inf_unemp.var2, impulse = "unemp.pre2", response = "infl.pre2", boot = T, n.ahead = 10,
ci=0.95))
```