Specification and Comparison of Seasonal ARIMA models on the Unemployment Rate in the USA

Author: Vincent Ibia

I. Purpose
Seasonal Autoregressive Integrated Moving Average models are a way to model univariate data that exhibit seasonality and cycles. Seasonality refers to a pattern in the data that repeats after a set number of time steps. The season itself id the point/time at which the pattern repeats. The pattern is also called the cycle.

II. Data
Our data comes from the (United States) Bureau of Labor Statistics. It's a simple record of the monthly unemployment rate. The data is not seasonally adjusted, since that's the goal of this exercise. Initially, I download the period from 1900-2018. There were many interesting things to observe in this series. However, the final timespan chosen to model was from December 2009 until December 2018.

<u>A. Initial Dataset 1990-2018</u>
The first thing we'll take a look at is the full dataset. In Graph 01, we have the raw time-series. We can see what looks like three large multi-year trends. One trough finishes around late 2000. This is the start of the dot com bust. The second finished in late 2007. This is the start of the Great Recession. The final dip is for the end of 2018, with a low of 3.5 seen in November that year. Unfortunately, the pattern implies the rate will be on the rise again.

Graph 01: Unadjusted unemployment rate in the United States 1990 – 2018

The Bureau of Labor Statistics holds data going back to 1948. In the Jan-1948 – Jan-2019 dataset, the minimum value is a rate of 2.4 percent, in October 1952. May 1969 was 2.9%, and the lowest since has been 3.5 percent in (October and) November 2018. Another implication that an increase could be on the horizon.

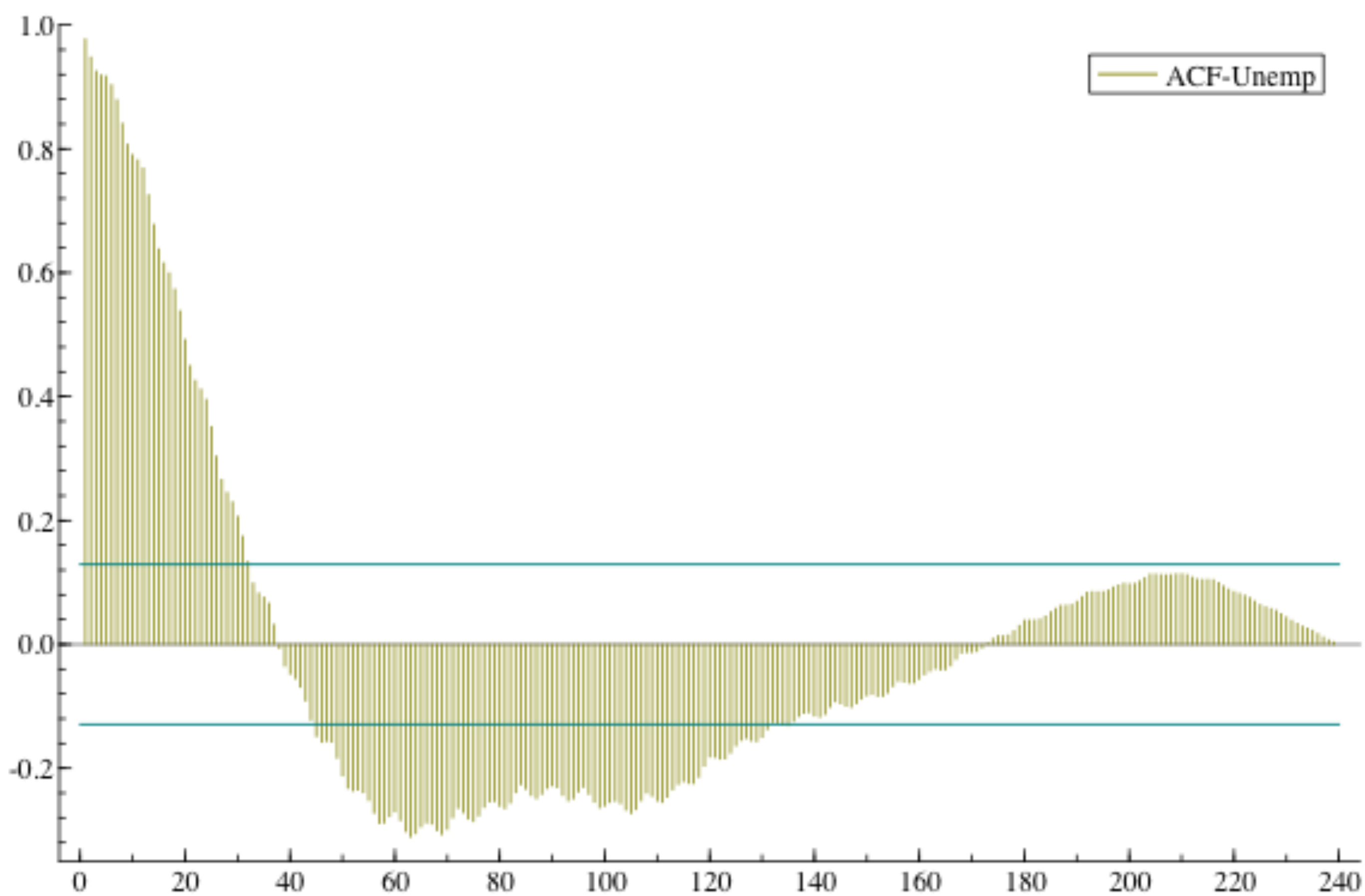
It's also useful to mention here that the data does not appear to be stationary. You could argue that it's somewhat mean-reverting (to 4%), but given the height and time span of the trend beyond 2008, that doesn't seem too accurate. Since the variance changes over time, the data is also heteroscedastic.

Our second graph (Graph 02) is the ACF of the unemployment rate. It uses a lag length equal to the full duration. We can see the autocorrelation between our final observation in 2018, and all months in previous years. There are very distinct "regime" changes, corresponding with what appeared to be our 3 large trends on Graph 01.

Regime 1: Dec-2018 to Dec-2015. Back towards Dec-2015, all unemployment rates have positive auto-correlation with Dec-2018.

Regime 2: Dec-2015 to Aug-2004. The observations here are all have negative auto-correlation with Regime 1.

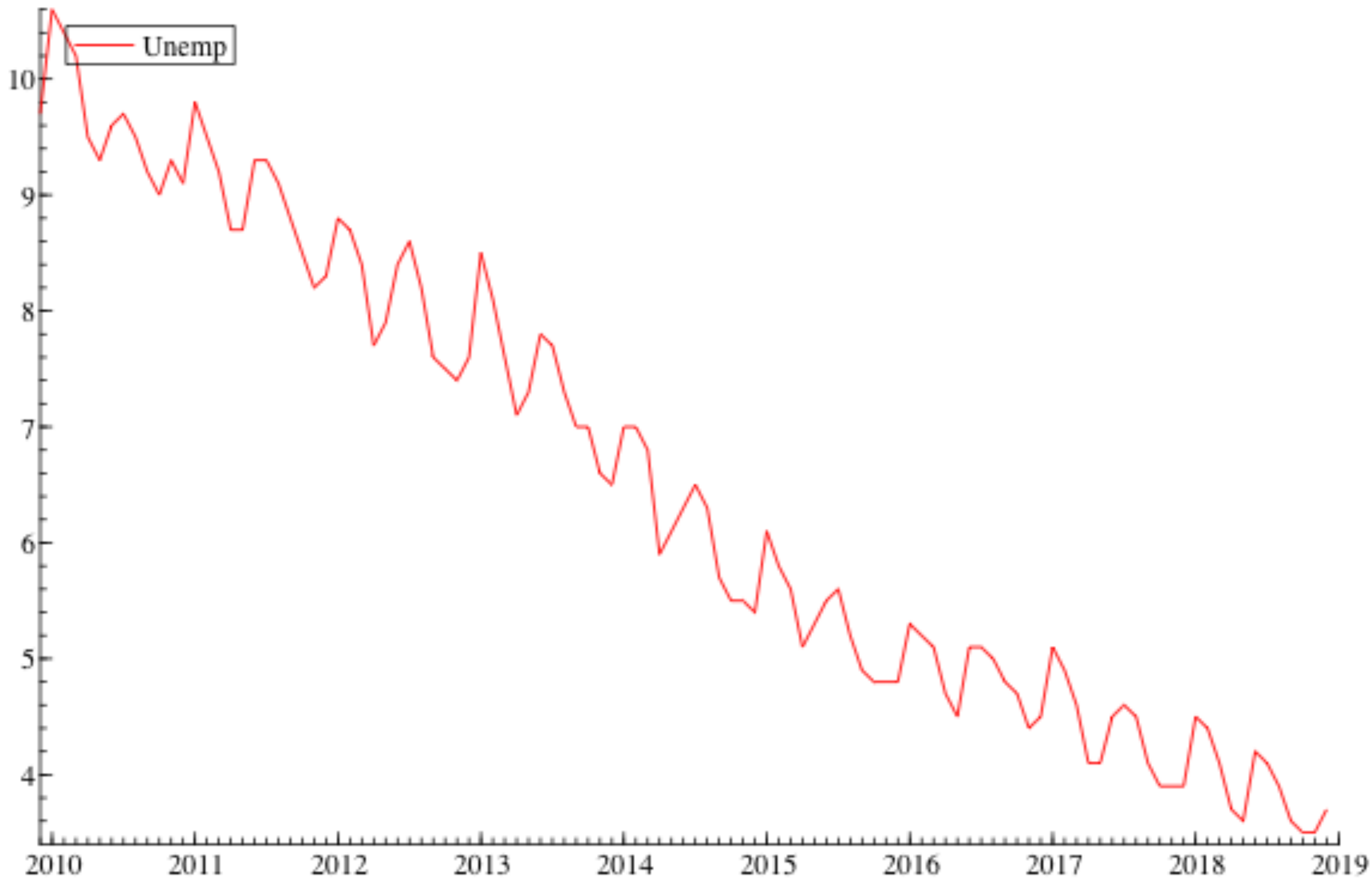
Regime 3: Aug-2004 to Jan-1999: This grouping correlates well with each other, and have less than statistically significant correlation with our Dec-2018 observation.



Graph 02: ACF of US unemployment rate 1990-2018
Shows periods of significant auto-correlation (positive or negative) with most recent data.

B. Truncated Dataset 2009-2018

Back in Graph 01, there is a clear downward trend in the data starting around mid-2009. Therefore, for simplicity, the set used for modeling will be from 2009-12 to 2018-12 (9 years). The shortened data is plotted in Graph 03.

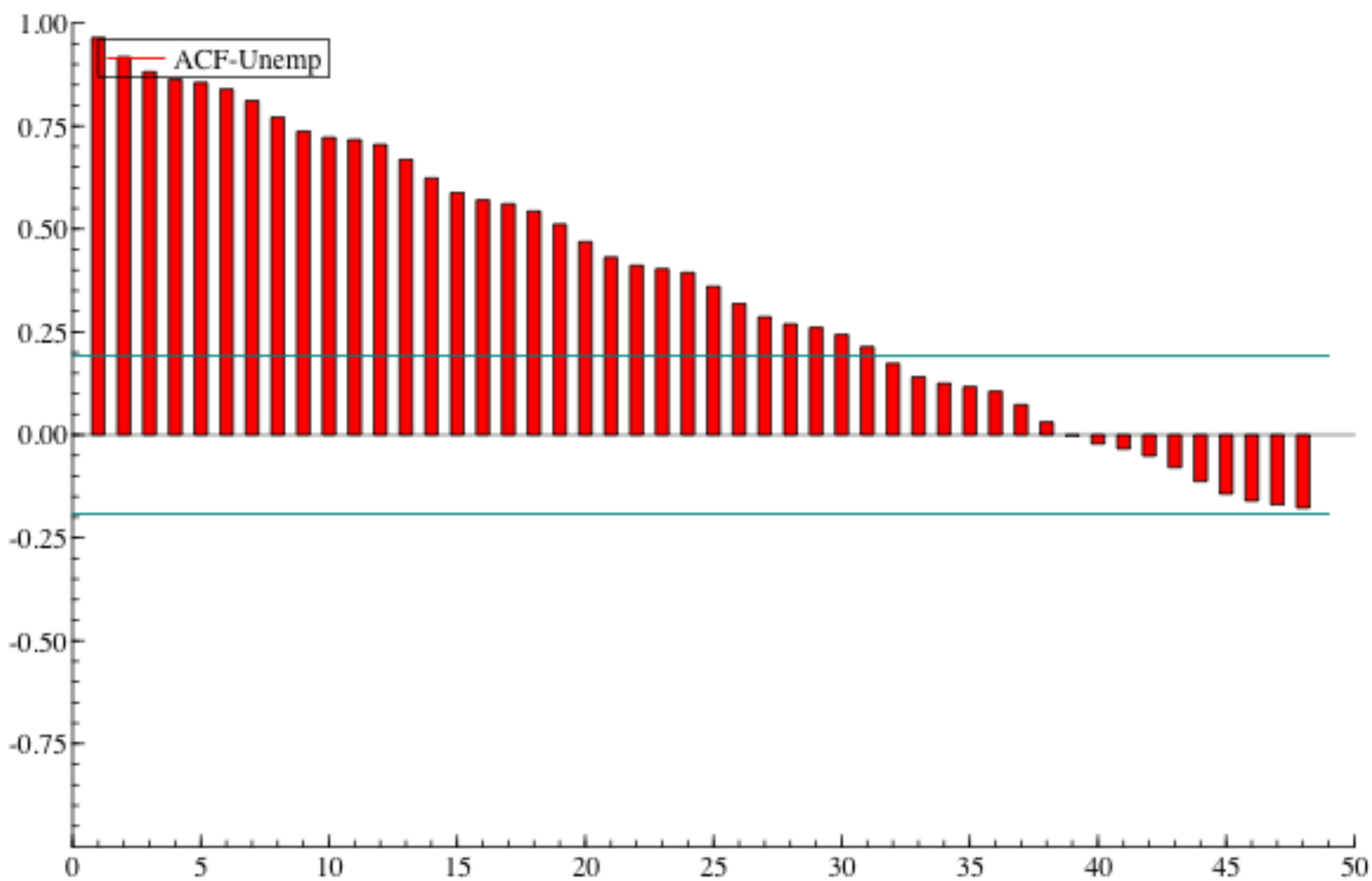


Graph 03: Nine-year trend of declining unemployment rate, with seasonal hikes

III. Determining the model Part I

A. Stationarity

The ACF plot of our 9-year period, is just a portion of what we saw in Graph 01. It doesn't help much. There's a slow decay, with statistical significance for 30 or so lags.



Graph 04: ACF of US unemployment rate, Dec-2009 to Dec-2018

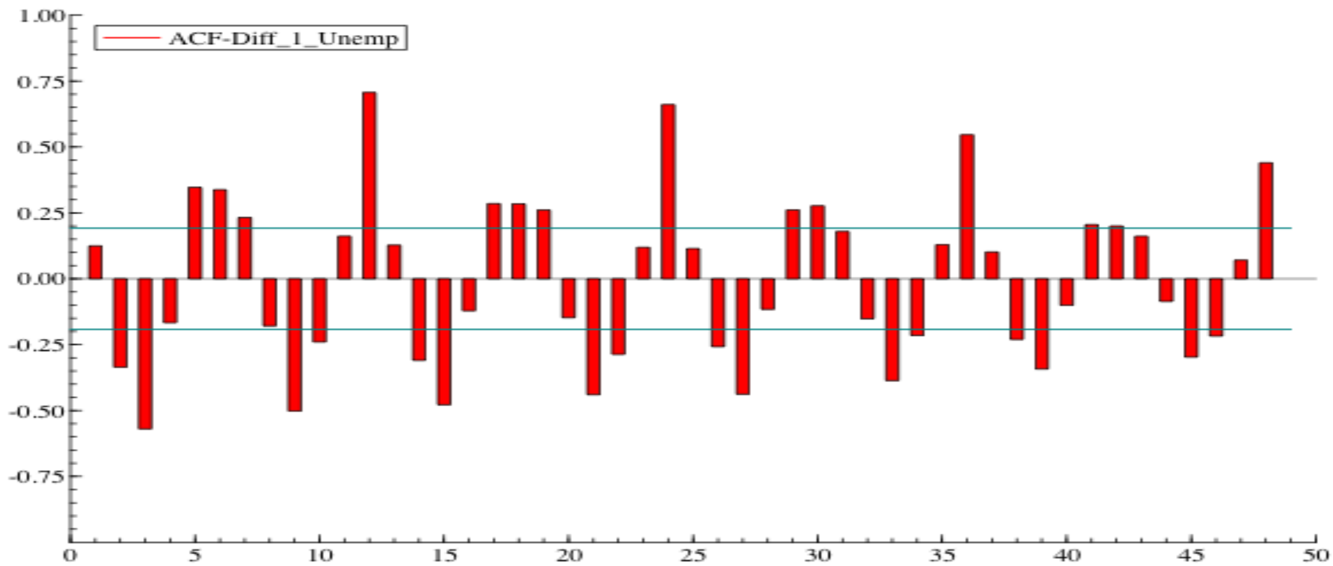
We'd like to figure out when the data is, or is most stationary. From the outset, our base data from 2009 on (Graph 03) is not stationary, but could arguably be called trend stationary. However, we will still give consideration to the first and second differences, to see if they pain a better picture. Moreover, these latter series will give us some intuition about our later model.

Our visual analysis of the autocorrelation functions didn't correspond well with initial beliefs. Moreover, the plots were somewhat misleading. What follows is a rapid description of these plots and assumed implications, so that we quickly get to the numerical analysis/results.

A1. Auto-correlation function plots

The time-series and ACF plots for both 1st and 2nd differences of the unemployment rate data are below. They exhibit some useful, some strange, and some misleading behavior.

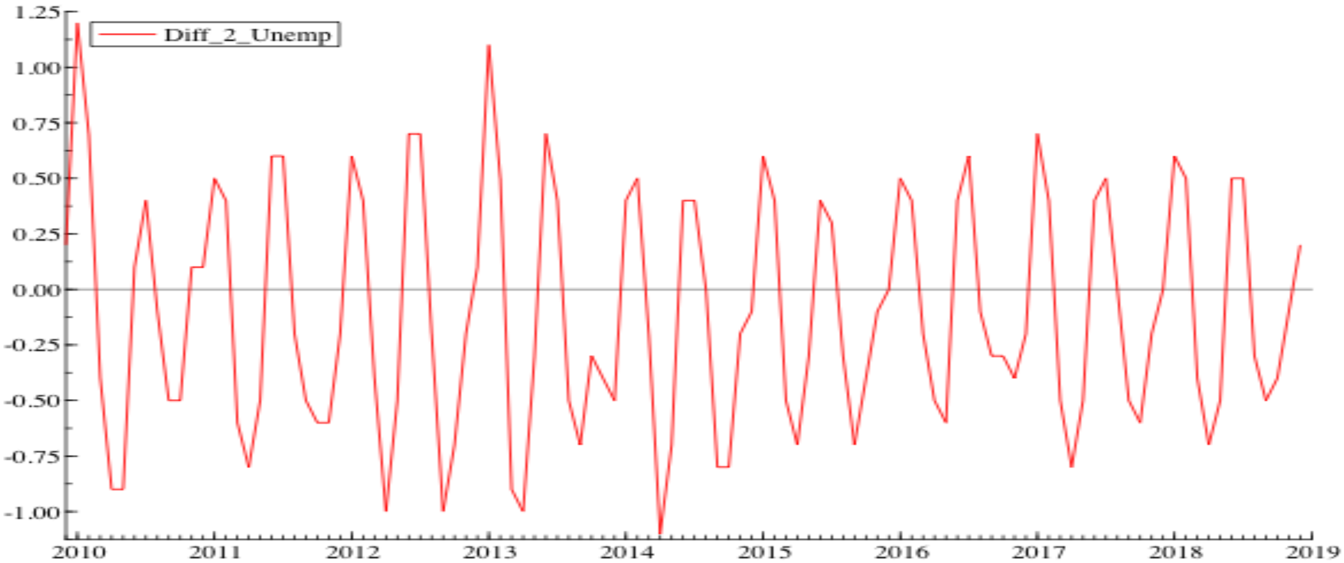
Graph 05a: 1st differenced unemployment rate



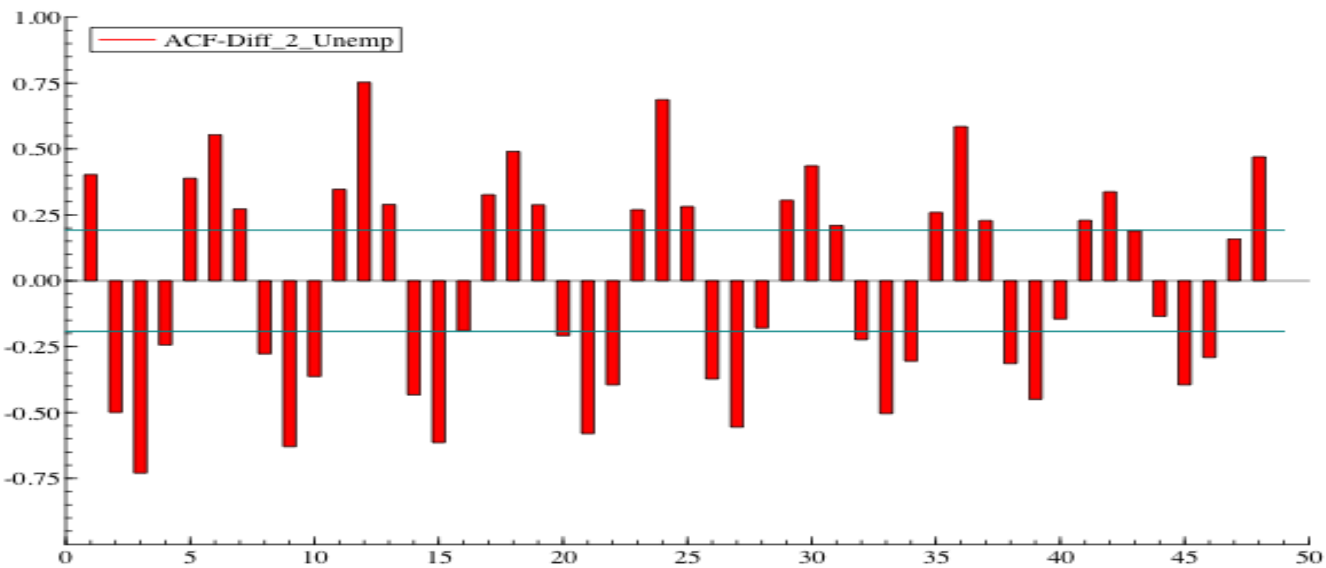
Graph 05b: ACF plot, 1st differenced unemployment rate

When using the 1st differences, our ACF for this 9-year period (Graph 05b), shows clear signs of yearly seasonality. There are prominent spikes in autocorrelation at multiples of 12 lags. Additionally, it shows what might be quarterly seasonality as well. In the differenced time-series (Graph 05a) we see a pattern that, if not centered around 0, seems to be stationary.

Switching to the 2nd differenced time-series. The time-series (Graph 06a) seems less stationary than the 1st differenced series. From 2009-2013, and then from 2015-2018 there seems to be slight trends upward. The plot does not appear to be centered at 0 either. Moving to the PACF of this series, we again see evidence of a 12-month season, but also potentially at 6 and 3 months. Moreover, the two months on either side a 3-month multiple are also statistically significant. It would not be better to model with this series.



Graph 06a: 2nd differenced unemployment rate



Graph 06b: ACF plot, 2nd differenced unemployment rate

A2. ADF Tests

Since our raw data has an obvious downward trend, we would want to run the ADF on this data assuming "trend", but can also check for trend and constant. On the other hand, we've calculated 1st and 2nd differences of the series. These series, visually, do not suggest a trend. Though there might, be a slight one that we cannot detect visually. The goal of our ADF testing is to find the setup that yields the lowest Information Criteria.

Table 01 shows the output of the Information Criteria for ADF test on each time-series. In bold, we can see, across all series, what the lowest results of the Information Criteria were. They all occur in the base time-series, under a test that assumes neither an intercept, nor a trend. This was unexpected.

For our **raw** data, considering 12 lags.

No Intercept No Trend

Information Criteria (to be minimized)

Akaike	**-0.472724**	Shibata	**-0.503884**
Schwarz	**-0.125469**	Hannan-Quinn	**-0.332358**

Intercept and No Trend

Information Criteria (to be minimized)

Akaike	-0.452532	Shibata	-0.488265
Schwarz	-0.078564	Hannan-Quinn	-0.301368

Intercept and Trend

Information Criteria (to be minimized)

Akaike	-0.432930	Shibata	-0.473496
Schwarz	-0.032250	Hannan-Quinn	-0.270969

For our **1st differenced** data, considering 12 lags.

No Intercept and no time trend

Information Criteria (to be minimized)

Akaike	-0.406170	Shibata	-0.437331
Schwarz	$\boxed{-0.058915}$	Hannan-Quinn	-0.265804

Intercept and No Trend

Information Criteria (to be minimized)

Akaike	$\boxed{-0.431907}$	Shibata	-0.467641
Schwarz	-0.057940	Hannan-Quinn	$\boxed{-0.280744}$

Intercept and Time Trend

Information Criteria (to be minimized)

Akaike	-0.428850	Shibata	$\boxed{-0.469416}$
Schwarz	-0.028171	Hannan-Quinn	-0.266889

For our **2nd differenced** data, considering 12 lags.

No Intercept no Trend

Information Criteria (to be minimized)

Akaike	-0.390189	Shibata	-0.421350
Schwarz	$\boxed{-0.042934}$	Hannan-Quinn	-0.249823

Intercept no Trend

Information Criteria (to be minimized)

Akaike	$\boxed{-0.404545}$	Shibata	-0.440278
Schwarz	-0.030577	Hannan-Quinn	$\boxed{-0.253381}$

Intercept and Trend

Information Criteria (to be minimized)

Akaike	-0.401276	Shibata	$\boxed{-0.441842}$
Schwarz	-0.000597	Hannan-Quinn	-0.239315

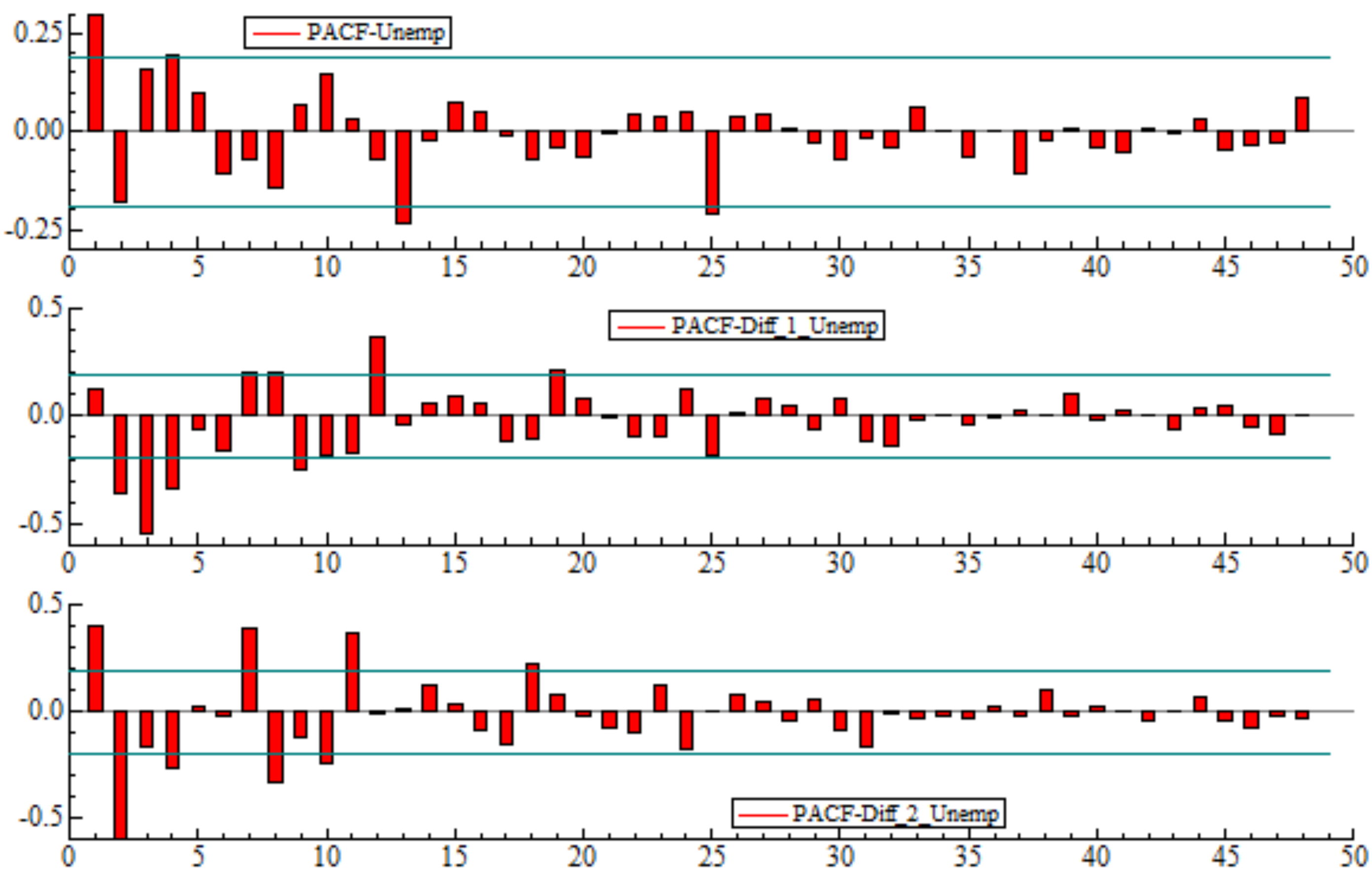
Table 01: Abbreviated ADF Test results

Not only was this the best results, but based on the ADF statistic and critical values for each of the 9 tests, only the raw data with no assumed intercept nor trend, was stationary. In every other test, the ADF statistic was not less than the critical value at any given alpha (1%, 5%, 10%). Full results for the ADF tests are in Appendix A. We will keep the results of the ADF test in mind when using the X12 Arima module on OxMetrics.

IV. Determining the Model Part 2

A. PACF plots

Non-stationary differenced data is not what we would expect. The PACF of each time-series will be used later on for some intuition on a decent model. For now, the plot for all three series is presented here (Graph 07), showing a lag of up to 48 months. Comparatively, the first plot behaves better. While the correlations taper off a bit faster in the last two plots, the last two plots are also quite complicated within the first 12 lags. Whereas the PACF of the raw unemployment rate, has seasonal spikes at lags 13 and 25; and, outside of lags 1 and 4, no other statistically significant lags prior to lag 13.



Graph 07: PACF. Unemployment rate (top), 1st difference (middle), 2nd difference (bottom)
Consecutively significant lags, and significant lags at irregular times (e.g. 7,11) are problematic.

B. Interpretations

There is strong evidence that differencing, either first or second is quite useless. By useless, I mean that after both 1st or 2nd differencing, our PACF indicates a series that is less stationary then the raw data. In fact, the PACF's for the differenced data implies non-stationarity.

In general, *the ACF controls the MA* - moving average order, and *the PACF controls the AR* - autoregressive order. Although both graphs help with guessing an ARMA specification. By "control", we mean that we can read the order from the graph, from the last point at which the lags are statistically significant.

This brings our attention back to Graph 07 above. The interpretations are:

- PACF raw
  Lag 2 is almost significant. Lag 4 is significant.
  Lag 13, and 25 are significant. However, 13+12 = 25
  We can say that this is a Seasonal effect.
  In short, we'd consider AR up to 4 lags.

- PACF 1st difference
  This is a mess. We get partial autocorrelation at lag 19, 12,9,8,7,4,3,2.
  How would we interpret this?
  You'll notice the ACF of this series is also quite interesting

- PACF 2nd difference
  Lags 18,11,10,8,7,4,2,1 are all significant.
  Lag 24 was on the border, but Lag 12 was near zero partial-autocorrelation
  I'd say an AR order of 18 is a bit excessive. There's no good pattern here.
  Maybe, the series is non-stationary

Our ADF tests confirmed our suspicions about Graph 07. The null hypotheses could not be rejected for neither the 1st differenced series, nor the 2nd differenced series. As such, I would think it strange to arbitrarily throw in 1st and 2nd differencing into consideration of our SARIMA model. We can look, but I expect to find nothing useful. Basically, a non-seasonal differencing parameter of $\boldsymbol{d = 0}$ should produce our best model.

V. Model Specification, Comparison, and Selection

A. Specification

There are 7 parameters for a Seasonal ARIMA model: 3 pairs of autoregressive, moving average, and integration parameters, and one final parameter to denote the length of the season.

$$\text{ARIMA}\ (p, d, q) * (P, D, Q)_S$$

The X12 Arima module in OXMetrics, as a setting for automatic model selection. In our previous discussion about the ACF and PACF plots, we decided that there was seasonality, with $S = 12$. It also made sense that the non-seasonal auto-regressive order of up to $p = 4$ could work. In total, 7 models were run, including one that used a log transformation, and a model generated my OxMetrics using the "Automatic ARIMA" option. The different specifications are listed below. In **bold** is the model I've selected. Full results for all of the model is in the Appendix.

1. Generic: S_{12}ARIMA (1,1,0) (1,1,0)
2. Generic: S_{12}ARIMA (0,1,0) (1,1,0)
3. 2nd differencing A: S_{12}ARIMA (0,2,0) (1,1,0)
4. 2nd differencing B: S_{12}ARIMA (3,2,0) (1,1,0)
5. Intuitive: S_{12}ARIMA (4,0,0) (2,1,0)
6. Auto-optimized: S_{12}ARIMA (3,1,1) (0,1,1)

For comparison, we will look at how forecasts from model #5 and #6 match up with the real data. Two quick notes on determining the models.

1. The "vanilla" (0,1,1) (0,1,1) model is the first assumption in automatic model selection. Since, OxMetrics did not consider it to be the best models, there's not real reason to compare it. However, I did compare the forecasts against Model-6 and consistently, each value was slightly greater than the forecasted points in Model-6.

2. The data implies an additive model: The seasonal pattern neither expands nor shrinks over time. When using the option for additive outliers, the results returned "No AO outliers identified". As such, none of the above listed models was made using *Outlier Detection* options.

B. Comparison

Model-6 is the best model according to the system. Compared to Model-5, it has higher Log likelihood, and low Information Criteria. There's also a lower variance, and std. err of variance. However, if we look into the result output for Model-6, it actually has issues with statistical significance of Non-Seasonal Lag 1 and Lag 2. Dividing Estimate by Standard Error, we get a value less than 2 is absolute value. Lag three is ok, but not by much. This lack of significance is also the problem with the Non-seasonal MA component. Additionally, AR Lag 1 Standard Error is larger than in Model-5. Essentially, these problems occur because of having $\boldsymbol{d} = 1$. Previously we saw that 1st differencing the data was a bit useless. Table 02 has Model-5 and Model-6 output for parameters.

Model-6 ARIMA Model: (3, 1, 1) (0, 1, 1)

Parameter	Estimate	Standard Errors
Nonseasonal AR		
Lag 1	.0345	.22571
Lag 2	-.1370	.09815
Lag 3	-.2284	.10202
Nonseasonal MA		
Lag 1	.3076	.22122
Seasonal MA		
Lag 12	.9998	.11226
Variance	.19780E-01	
SE of Var	.30521E-02	

Model-5 ARIMA Model: (4, 0, 0) (2, 1, 0)

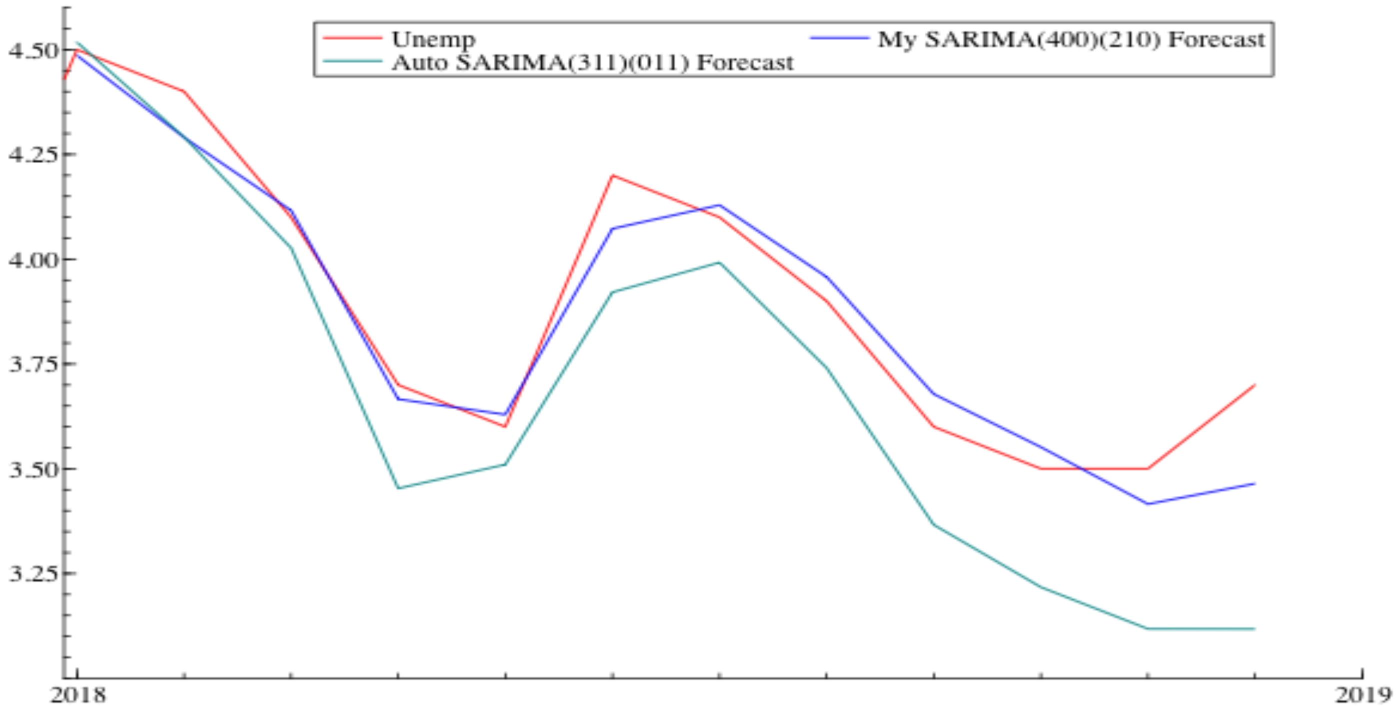
Parameter	Estimate	Standard Errors
Nonseasonal AR		
Lag 1	.7399	.10279
Lag 2	.0732	.12824
Lag 3	-.0569	.12626
Lag 4	.2389	.10399
Seasonal AR		
Lag 12	-.7079	.10691
Lag 24	-.2459	.10940
Variance	.27893E-01	
SE of Var	.42786E-02	

Table 02: Estimated parameter comparison, Model-6 and Model-5

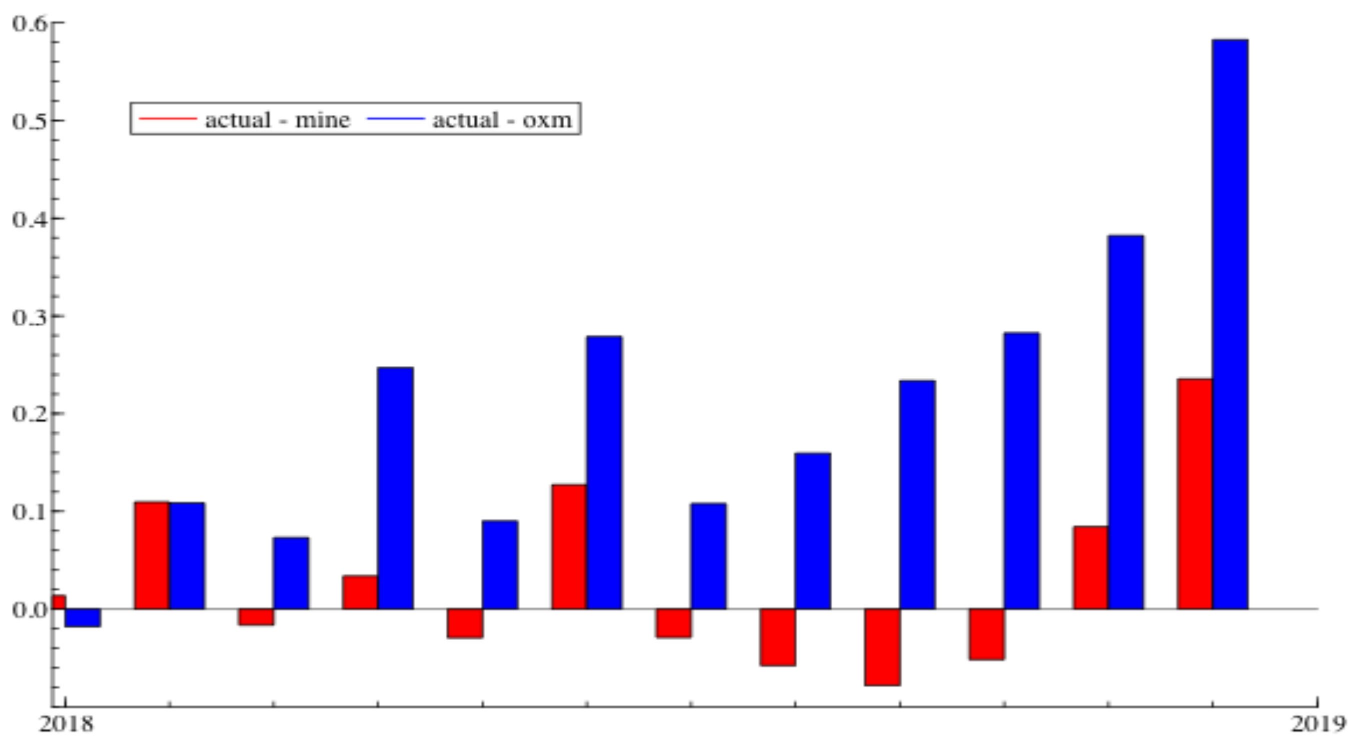
C. Selection

While the Likelihood Statistics declare Model-5 the winner, I disagree. The parameters used don't make sense. If we look back at Graph 05b, ACF of 1st differences, and Graph 07 of the PACF's, what exactly motivates the reason to have $p = 3$ and $q = 1$? Nothing really. On the other hand, we discussed sound reasons for the Model-5 parameters. However, since the ACF of the raw series (Graph 04) wasn't too nice, admittedly there was no goo way to decide on the MA order. Something like 30 would be too many lags to include.

Visually, my select model, Model-5 is a clear winner, based on the forecast. Model-6 plays it safe. For the entire forecast is underestimates the unemployment rate. Realistically, that's bad. Moreover, in general it's further away from the actual values (except for the 1st estimate), than Model-5. It seems Model-5 is penalized since it slightly overestimates the unemployment are for various points. Graph 08 shows the forecasts plotted with the actual values. Graph 08 is a simple bar plot of (Actual Value – Model Forecasted Value) to further illustrate the reason for my selection.



Graph 08: Actual Unemployment rate, and Forecasts: My/Model-5, Auto/Model-6



Graph 09: Forecast Error, red – Model-5 (Mine), blue – Model-6 (OxMetrics)

To briefly conclude, while Model-6 did have better results for matching the sample data, my selected Model-5 was superior in the forecast, and would be the one to go with.

VI. Appendix

A. Full output for Augmented Dickey-Fuller tests

A1. ADF Test, Unemployment rate, base data

```
TESTS :
---------------------------------------------
ADF Test with 12 lags
No intercept and no time trend
H0: Unemp is I(1)

ADF Statistics: -2.59152
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)
              1%        5%       10%
         -2.56572  -1.94093  -1.61663

OLS Results
             Coefficient    t-value
y_1            -0.018579    -2.5915
dy_1           -0.350456    -3.6357
dy_2           -0.338199    -3.3752
dy_3           -0.384948    -3.6617
dy_4           -0.217074    -1.9255
dy_5            0.045419    0.39564
dy_6           -0.021181   -0.18361
dy_7            0.078827    0.68828
dy_8            0.113341    0.99000
dy_9           -0.079062   -0.70322
dy_10          -0.160711    -1.5420
dy_11          -0.111135    -1.0967
dy_12           0.401754     4.3490
RSS             2.645402
OBS            96.000000
Information Criteria (to be minimized)
Akaike         -0.472724  Shibata          -0.503884
Schwarz        -0.125469  Hannan-Quinn     -0.332358

TESTS :
---------------------------------------------
ADF Test with 12 lags
Intercept and no time trend
H0: Unemp is I(1)

ADF Statistics: -1.37578
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)
              1%        5%       10%
          -3.4323  -2.86228  -2.56721

OLS Results
             Coefficient    t-value
y_1            -0.014992    -1.3758
dy_1           -0.360416    -3.6227
```

dy_2	-0.350053	-3.3577
dy_3	-0.398539	-3.6202
dy_4	-0.231399	-1.9627
dy_5	0.030670	0.25525
dy_6	-0.034920	-0.29081
dy_7	0.065913	0.55486
dy_8	0.102076	0.86596
dy_9	-0.087828	-0.76551
dy_10	-0.167509	-1.5822
dy_11	-0.116883	-1.1384
dy_12	0.394923	4.1956
Constant	-0.031933	-0.43884
RSS	2.639203	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.452532	Shibata	-0.488265
Schwarz	-0.078564	Hannan-Quinn	-0.301368

TESTS :

ADF Test with 12 lags
Intercept and time trend
H0: Unemp is I(1)

ADF Statistics: -0.675754
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-3.96104	-3.41127	-3.12748

OLS Results

	Coefficient	t-value
y_1	-0.056336	-0.67575
dy_1	-0.318105	-2.4296
dy_2	-0.311039	-2.3820
dy_3	-0.363116	-2.7651
dy_4	-0.199370	-1.4808
dy_5	0.063472	0.46205
dy_6	-0.001222	-0.0088470
dy_7	0.097626	0.72246
dy_8	0.134413	0.99630
dy_9	-0.058246	-0.44963
dy_10	-0.144612	-1.2489
dy_11	-0.098970	-0.90645
dy_12	0.406073	4.1798
Constant	0.243573	0.43843
Trend	-0.002750	-0.50026
RSS	2.631074	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.432930	Shibata	-0.473496
Schwarz	-0.032250	Hannan-Quinn	-0.270969

A2. ADF Test, 1st differences of Unemployment rate

TESTS :

ADF Test with 12 lags
No intercept and no time trend
H0: Diff_1_Unemp is I(1)

ADF Statistics: -1.02122
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-2.56572	-1.94093	-1.61663

OLS Results

	Coefficient	t-value
y_1	-0.300457	-1.0212
dy_1	-1.012798	-3.3453
dy_2	-1.250332	-4.0815
dy_3	-1.518229	-4.8091
dy_4	-1.592918	-4.7660
dy_5	-1.399044	-3.9490
dy_6	-1.263128	-3.5638
dy_7	-1.013631	-2.9774
dy_8	-0.736959	-2.3734
dy_9	-0.649565	-2.5285
dy_10	-0.647637	-3.2961
dy_11	-0.616194	-4.2783
dy_12	-0.095734	-0.96922
RSS	2.827454	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.406170	Shibata	-0.437331
Schwarz	-0.058915	Hannan-Quinn	-0.265804

TESTS :

ADF Test with 12 lags
Intercept and no time trend
H0: Diff_1_Unemp is I(1)

ADF Statistics: -2.25524
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-3.4323	-2.86228	-2.56721

OLS Results

	Coefficient	t-value
y_1	-1.770243	-2.2552
dy_1	0.399206	0.52418
dy_2	0.061829	0.086153
dy_3	-0.321524	-0.47970
dy_4	-0.529396	-0.85140
dy_5	-0.474992	-0.82486

dy_6	-0.480121	-0.92002
dy_7	-0.376743	-0.81862
dy_8	-0.236780	-0.60205
dy_9	-0.278672	-0.89213
dy_10	-0.396337	-1.7247
dy_11	-0.469576	-2.9517
dy_12	-0.042677	-0.42457
Constant	-0.101437	-2.0139
RSS	2.694201	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.431907	Shibata	-0.467641
Schwarz	-0.057940	Hannan-Quinn	-0.280744

TESTS :

ADF Test with 12 lags
Intercept and time trend
H0: Diff_1_Unemp is I(1)

ADF Statistics: -2.57428
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-3.96104	-3.41127	-3.12748

OLS Results

	Coefficient	t-value
y_1	-2.173735	-2.5743
dy_1	0.783221	0.95858
dy_2	0.422973	0.54950
dy_3	0.016064	0.022343
dy_4	-0.220430	-0.33109
dy_5	-0.201247	-0.32827
dy_6	-0.244675	-0.44314
dy_7	-0.182460	-0.37739
dy_8	-0.086386	-0.21099
dy_9	-0.171578	-0.53204
dy_10	-0.328503	-1.3971
dy_11	-0.438096	-2.7307
dy_12	-0.035465	-0.35355
Constant	-0.125208	-2.3370
Trend	0.000918	1.2672
RSS	2.641830	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.428850	Shibata	-0.469416
Schwarz	-0.028171	Hannan-Quinn	-0.266889

A3. ADF Test, 2nd differences of Unemployment rate

TESTS :

ADF Test with 12 lags
No intercept and no time trend
H0: Diff_2_Unemp is I(1)

ADF Statistics: -0.899516
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-2.56572	-1.94093	-1.61663

OLS Results

	Coefficient	t-value
y_1	-0.133898	-0.89952
dy_1	-0.228020	-1.2994
dy_2	-1.163495	-6.7481
dy_3	-0.547854	-2.8433
dy_4	-1.224532	-6.2468
dy_5	-0.390890	-1.8521
dy_6	-1.081098	-5.1140
dy_7	-0.194561	-0.97044
dy_8	-0.764772	-3.9291
dy_9	-0.151056	-0.95956
dy_10	-0.715038	-4.9028
dy_11	-0.111763	-1.0962
dy_12	-0.142546	-1.4345
RSS	2.873003	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.390189	Shibata	-0.421350
Schwarz	-0.042934	Hannan-Quinn	-0.249823

TESTS :

ADF Test with 12 lags
Intercept and no time trend
H0: Diff_2_Unemp is I(1)

ADF Statistics: -1.96477
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-3.4323	-2.86228	-2.56721

OLS Results

	Coefficient	t-value
y_1	-0.810843	-1.9648
dy_1	0.401915	1.0086
dy_2	-0.550203	-1.4157
dy_3	-0.018181	-0.050965

dy_4	-0.727988	-2.1238
dy_5	0.015211	0.048849
dy_6	-0.702305	-2.3390
dy_7	0.086287	0.33895
dy_8	-0.509190	-2.1115
dy_9	0.016780	0.091933
dy_10	-0.574866	-3.4904
dy_11	-0.053048	-0.49993
dy_12	-0.096995	-0.95547
Constant	-0.092627	-1.7555
RSS	2.768939	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.404545	Shibata	-0.440278
Schwarz	-0.030577	Hannan-Quinn	-0.253381

TESTS :

ADF Test with 12 lags
Intercept and time trend
H0: Diff_2_Unemp is I(1)

ADF Statistics: -2.30016
Asymptotic critical values, Davidson, R. and MacKinnon, J. (1993)

1%	5%	10%
-3.96104	-3.41127	-3.12748

OLS Results

	Coefficient	t-value
y_1	-1.020002	-2.3002
dy_1	0.592292	1.3942
dy_2	-0.361769	-0.87147
dy_3	0.148931	0.39256
dy_4	-0.569207	-1.5636
dy_5	0.146768	0.44832
dy_6	-0.582443	-1.8553
dy_7	0.176847	0.67078
dy_8	-0.435091	-1.7588
dy_9	0.063915	0.34422
dy_10	-0.542450	-3.2655
dy_11	-0.044401	-0.41905
dy_12	-0.091249	-0.90117
Constant	-0.117287	-2.0907
Trend	0.000926	1.2603
RSS	2.715690	
OBS	96.000000	

Information Criteria (to be minimized)

Akaike	-0.401276	Shibata	-0.441842
Schwarz	-0.000597	Hannan-Quinn	-0.239315

B. Full results for SARIMA models

B1. S_{12}ARIMA (1,1,0) (1,1,0)

```
MODEL DEFINITION for Unemp
 Transformation:      No transformation
 ARIMA Model:  (1 1 0)(1 1 0)
 regARIMA Model Span: 2009.Dec to 2017.Dec

 MODEL ESTIMATION/EVALUATION
 Estimation converged in    5 ARMA iterations,   16 function evaluations.
    ARIMA Model:  (1 1 0)(1 1 0)
         Nonseasonal differences: 1
         Seasonal differences:    1
Parameter                    Estimate         Standard Errors
 -----------------------------------------------------
 Nonseasonal AR
   Lag  1                       -.1999          .10544
 Seasonal AR
   Lag 12                       -.5831          .08458
 Variance                        .31957E-01
 SE of Var                       .49311E-02
 -----------------------------------------------------

 Likelihood Statistics
 ------------------------------------------------------------------
 Number of observations (nobs)                                  97
 Effective number of observations (nefobs)                      84
 Number of parameters estimated (np)                             3
 Log likelihood (L)                                        22.9167
 AIC                                                      -39.8334
 AICC (F-corrected-AIC)                                   -39.5334
 Hannan Quinn                                             -36.9019
 BIC                                                      -32.5410
 ------------------------------------------------------------------
 Roots of ARIMA Model
  Root                  Real  Imaginary    Modulus  Frequency
  ------------------------------------------------------------
  Nonseasonal AR
   Root  1           -5.0024      .0000     5.0024      .5000
  Seasonal AR
   Root  1           -1.7150      .0000     1.7150      .5000
  ------------------------------------------------------------
 FORECASTING
  Origin  2017.Dec
  Number        12

  Confidence intervals with coverage probability ( .95000)
   ---------------------------------------
       Date      Lower  Forecast     Upper
   ---------------------------------------
   2018.Jan       4.10      4.45      4.80
   2018.Feb       3.86      4.31      4.76
   2018.Mar       3.59      4.12      4.66
   2018.Apr       3.07      3.68      4.29
   2018.May       2.89      3.57      4.24
   2018.Jun       3.34      4.08      4.82
   2018.Jul       3.33      4.12      4.92
   2018.Aug       3.18      4.02      4.87
   2018.Sep       2.85      3.74      4.63
   2018.Oct       2.66      3.60      4.54
   2018.Nov       2.44      3.42      4.41
   2018.Dec       2.45      3.48      4.51
```

B2. S_{12}ARIMA (0,1,0) (1,1,0)

```
MODEL DEFINITION for Unemp
 Transformation:      No transformation
 ARIMA Model:  (0 1 0)(1 1 0)
 regARIMA Model Span: 2009.Dec to 2017.Dec


 MODEL ESTIMATION/EVALUATION
 Estimation converged in    4 ARMA iterations,    9 function evaluations.
 ARIMA Model:  (0 1 0)(1 1 0)
   Nonseasonal differences: 1
   Seasonal differences:    1
                                                 Standard
 Parameter                     Estimate            Errors
 -----------------------------------------------------
 Seasonal AR
   Lag 12                        -.5783            .08550

 Variance                           .33308E-01
 SE of Var                          .51396E-02
 -----------------------------------------------------

 Likelihood Statistics
 ------------------------------------------------------------------
 Number of observations (nobs)                                   97
 Effective number of observations (nefobs)                       84
 Number of parameters estimated (np)                              2
 Log likelihood (L)                                         21.2479
 AIC                                                       -38.4957
 AICC (F-corrected-AIC)                                    -38.3476
 Hannan Quinn                                              -36.5414
 BIC                                                       -33.6341
 ------------------------------------------------------------------

 Roots of ARIMA Model
  Root                   Real  Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  Seasonal AR
   Root  1           -1.7291      .0000     1.7291      .5000
  -----------------------------------------------------------


 FORECASTING
  Origin  2017.Dec
  Number        12

  Confidence intervals with coverage probability ( .95000)
   ---------------------------------------
        Date      Lower  Forecast     Upper
   ---------------------------------------
   2018.Jan       4.08      4.44      4.80
   2018.Feb       3.79      4.30      4.81
   2018.Mar       3.50      4.12      4.74
   2018.Apr       2.96      3.67      4.39
   2018.May       2.76      3.56      4.36
   2018.Jun       3.20      4.07      4.95
   2018.Jul       3.17      4.12      5.06
   2018.Aug       3.00      4.02      5.03
   2018.Sep       2.66      3.73      4.80
   2018.Oct       2.46      3.59      4.72
   2018.Nov       2.23      3.42      4.60
   2018.Dec       2.23      3.47      4.71
   ---------------------------------------
```

B3. S_{12}ARIMA (0,2,0) (1,1,0)

```
MODEL DEFINITION for Unemp
 Transformation:      No transformation
 ARIMA Model:  (0 2 0)(1 1 0)
 regARIMA Model Span: 2009.Dec to 2017.Dec


 MODEL ESTIMATION/EVALUATION
 Estimation converged in    4 ARMA iterations,    9 function evaluations.
 ARIMA Model:  (0 2 0)(1 1 0)
   Nonseasonal differences: 2
   Seasonal differences:    1
                                              Standard
 Parameter                   Estimate          Errors
 -----------------------------------------------------
 Seasonal AR
   Lag 12                      -.5758          .08653

 Variance                        .79383E-01
 SE of Var                       .12323E-01
 -----------------------------------------------------

 Likelihood Statistics
 ------------------------------------------------------------------
 Number of observations (nobs)                                   97
 Effective number of observations (nefobs)                       83
 Number of parameters estimated (np)                              2
 Log likelihood (L)                                        -15.0496
 AIC                                                        34.0993
 AICC (F-corrected-AIC)                                     34.2493
 Hannan Quinn                                               36.0428
 BIC                                                        38.9370
 ------------------------------------------------------------------

 Roots of ARIMA Model
  Root                  Real  Imaginary    Modulus  Frequency
  -----------------------------------------------------------
  Seasonal AR
   Root  1           -1.7367      .0000     1.7367      .5000
  -----------------------------------------------------------


 FORECASTING
  Origin  2017.Dec
  Number        12

  Confidence intervals with coverage probability ( .95000)
   ---------------------------------------
        Date      Lower  Forecast     Upper
   ---------------------------------------
   2018.Jan      3.848     4.400     4.952
   2018.Feb      2.980     4.215     5.450
   2018.Mar      1.922     3.988     6.054
   2018.Apr       .478     3.503     6.528
   2018.May      -.750     3.345     7.441
   2018.Jun     -1.450     3.818     9.086
   2018.Jul     -2.716     3.818    10.352
   2018.Aug     -4.211     3.676    11.563
   2018.Sep     -5.974     3.349    12.671
   2018.Oct     -7.672     3.164    13.999
   2018.Nov     -9.473     2.949    15.370
   2018.Dec    -11.115     2.964    17.043
   ---------------------------------------
```

B4. S_{12}ARIMA (3,2,0) (1,1,0)

```
MODEL DEFINITION for Unemp
 Transformation:  No transformation    ARIMA Model:  (3 2 0)(1 1 0)
 regARIMA Model Span: 2009.Dec to 2017.Dec

 MODEL ESTIMATION/EVALUATION
 Estimation converged in    6 ARMA iterations,   31 function evaluations.
 ARIMA Model:  (3 2 0)(1 1 0)
      Nonseasonal differences: 2
      Seasonal differences:    1
Parameter                          Estimate       Standard Errors
 -----------------------------------------------------
 Nonseasonal AR
      Lag  1                     -.8126          .10192
      Lag  2                     -.5376          .12463
      Lag  3                     -.3256          .10295
 Seasonal AR
      Lag 12                     -.5885          .08506
 Variance                      .45044E-01
 SE of Var                     .69922E-02
 -----------------------------------------------------
 Likelihood Statistics
 ------------------------------------------------------------------
 Number of observations (nobs)                                   97
 Effective number of observations (nefobs)                       83
 Number of parameters estimated (np)                              5
 Log likelihood (L)                                          7.9039
 AIC                                                        -5.8078
 AICC (F-corrected-AIC)                                     -5.0286
 Hannan Quinn                                                -.9490
 BIC                                                         6.2864
 ------------------------------------------------------------------
 Roots of ARIMA Model
  Root                  Real  Imaginary    Modulus  Frequency
  ------------------------------------------------------------
  Nonseasonal AR
   Root  1            -.1165     1.4670     1.4716      .2626
   Root  2            -.1165    -1.4670     1.4716     -.2626
   Root  3           -1.4184      .0000     1.4184      .5000
  Seasonal AR
   Root  1           -1.6992      .0000     1.6992      .5000
  ------------------------------------------------------------

 FORECASTING
     Origin  2017.Dec
     Number        12
  Confidence intervals with coverage probability ( .95000)
   ---------------------------------------
       Date      Lower  Forecast     Upper
   ---------------------------------------
   2018.Jan      3.984     4.400     4.816
   2018.Feb      3.588     4.234     4.879
   2018.Mar      3.170     4.067     4.964
   2018.Apr      2.426     3.600     4.773
   2018.May      1.940     3.463     4.985
   2018.Jun      2.087     3.964     5.842
   2018.Jul      1.740     3.997     6.253
   2018.Aug      1.220     3.878     6.536
   2018.Sep       .493     3.580     6.668
   2018.Oct      -.110     3.424     6.959
   2018.Nov      -.768     3.234     7.235
   2018.Dec     -1.211     3.276     7.764
   ---------------------------------------
```

B5. S_{12}ARIMA (4,0,0) (2,1,0) *Author's Selected Model*

```
MODEL DEFINITION for Unemp
   Transformation: No transformation
   ARIMA Model:  (4 0 0)(2 1 0)
   regARIMA Model Span: 2009.Dec to 2017.Dec

MODEL ESTIMATION/EVALUATION
   Estimation converged in 16 ARMA iterations, 115 function evaluations.
      ARIMA Model: (4 0 0)(2 1 0)         Seasonal differences:    1
Parameter                        Estimate         Standard Errors
 -----------------------------------------------------
 Nonseasonal AR
      Lag  1                       .7399          .10279
      Lag  2                       .0732          .12824
      Lag  3                      -.0569          .12626
      Lag  4                       .2389          .10399
 Seasonal AR
      Lag 12                      -.7079          .10691
      Lag 24                      -.2459          .10940
 Variance                    .27893E-01
 SE of Var                   .42786E-02
 -----------------------------------------------------
 Likelihood Statistics
 ------------------------------------------------------------------
 Number of observations (nobs)                                   97
 Effective number of observations (nefobs)                       85
 Number of parameters estimated (np)                              7
 Log likelihood (L)                                         26.8876
 AIC                                                       -39.7752
 AICC (F-corrected-AIC)                                    -38.3207
 Hannan Quinn                                              -32.8977
 BIC                                                       -22.6767
 ------------------------------------------------------------------
 Roots of ARIMA Model
       Root              Real      Imaginary  Modulus    Frequency
  -----------------------------------------------------------
  Nonseasonal AR
      Root  1           1.0029      .0000     1.0029      .0000
      Root  2          -1.6489      .0000     1.6489      .5000
      Root  3            .4420     1.5283     1.5909      .2052
      Root  4            .4420    -1.5283     1.5909     -.2052
  Seasonal AR
      Root  1          -1.4391     1.4123     2.0164      .3765
      Root  2          -1.4391    -1.4123     2.0164     -.3765
  -----------------------------------------------------------
 FORECASTING:  Origin  2017.Dec  Number       12
  Confidence intervals with coverage probability ( .95000)
   ---------------------------------------
       Date      Lower  Forecast     Upper
   ---------------------------------------
   2018.Jan       4.16      4.49      4.81
   2018.Feb       3.88      4.29      4.70
   2018.Mar       3.66      4.12      4.57
   2018.Apr       3.19      3.67      4.15
   2018.May       3.11      3.63      4.14
   2018.Jun       3.52      4.07      4.62
   2018.Jul       3.54      4.13      4.72
   2018.Aug       3.34      3.96      4.57
   2018.Sep       3.03      3.68      4.32
   2018.Oct       2.88      3.55      4.22
   2018.Nov       2.72      3.42      4.11
   2018.Dec       2.74      3.46      4.19
   ---------------------------------------
```

B6. S_{12}ARIMA (3,1,1) (0,1,1) *OxMetrics' Selected Model*

```
     Series Title- X-12-ARIMA run of Unemp
     Series Name- Unemp

        -Period covered- 12th month,2009 to 12th month,2017
                        Automatic ARIMA Model Selection

                    Procedure based closely on TRAMO method
                         of Gomez and Maravall (2000)
               "Automatic Modeling Methods for Univariate Series",
                            A Course in Time Series
                   (Edited by D. Pena, G. C. Tiao, R. S. Tsay),
                         New York : J. Wiley and Sons


     Maximum order for regular ARMA parameters :   2
     Maximum order for seasonal ARMA parameters :   1
     Maximum order for regular differencing :   2
     Maximum order for seasonal differencing :   1

   Results of Unit Root Test for identifying orders of differencing:
     Regular difference order :   1
     Seasonal difference order :   1

     Mean is not significant.

  Automatic model choice : (0 1 1)(0 1 1)

  Final automatic model choice : (3 1 1)(0 1 1)
   End of automatic model selection procedure.
 Estimation converged in   10 ARMA iterations,   61 function evaluations.
 ARIMA Model:  (3 1 1)(0 1 1)
   Nonseasonal differences: 1
   Seasonal differences:    1
                                                 Standard
 Parameter                    Estimate            Errors
 -----------------------------------------------------
 Nonseasonal AR
   Lag  1                        .0345            .22571
   Lag  2                       -.1370            .09815
   Lag  3                       -.2284            .10202

 Nonseasonal MA
   Lag  1                        .3076            .22122

 Seasonal MA
   Lag 12                        .9998            .11226

 Variance                        .19780E-01
 SE of Var                       .30521E-02
 -----------------------------------------------------
```

Likelihood Statistics

Number of observations (nobs)	97
Effective number of observations (nefobs)	84
Number of parameters estimated (np)	6
Log likelihood (L)	32.9541
AIC	-53.9083
AICC (F-corrected-AIC)	-52.8174
Hannan Quinn	-48.0453
BIC	-39.3234

Roots of ARIMA Model

Root	Real	Imaginary	Modulus	Frequency
Nonseasonal AR				
Root 1	.6484	1.3740	1.5193	.1798
Root 2	.6484	-1.3740	1.5193	-.1798
Root 3	-1.8964	.0000	1.8964	.5000
Nonseasonal MA				
Root 1	3.2510	.0000	3.2510	.0000
Seasonal MA				
Root 1	1.0002	.0000	1.0002	.0000

FORECASTING
Origin 2017.Dec
Number 12

Confidence intervals with coverage probability (.95000)

Date	Lower	Forecast	Upper
2018.Jan	4.24	4.52	4.79
2018.Feb	3.95	4.29	4.63
2018.Mar	3.65	4.03	4.40
2018.Apr	3.06	3.45	3.84
2018.May	3.10	3.51	3.92
2018.Jun	3.49	3.92	4.36
2018.Jul	3.53	3.99	4.45
2018.Aug	3.26	3.74	4.23
2018.Sep	2.86	3.37	3.87
2018.Oct	2.69	3.22	3.74
2018.Nov	2.57	3.12	3.66
2018.Dec	2.56	3.12	3.68

VII. References

United States Department of Labor - Bureau of Labor Statistics
https://beta.bls.gov/dataViewer/view/timeseries/LNU04000000