**Time Series Analysis and Forecasting**

**Fernando Oliveira**
Oliveira@essec.edu

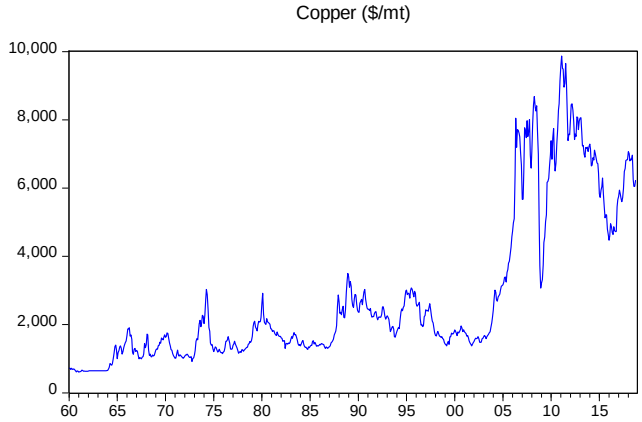
**Workshop 4: Stationarity and ARIMA**

This workshop consists of two main sections:
(a) Dickey-Fuller Tests
(b) Arima Modelling

## (a) Dickey-Fuller Tests

In this section, we shall test for stationarity in a time series copper prices from 1960-Jan to 2018-Dec. The series for this first part in stored stored in the EViews file **session 4 workshop.wfl** and we want to analyse the series of **copper** prices.

(i) First plot the series by selecting from the main menu **Quick** **Graph**




Plot the correlogram by selecting **Quick** **Series Statistics** **Correlogram**

Specify **copper**, then click **OK**

Does the series look stationary?

(ii) Carry out a basic Dickey-Fuller test with no constant or trend term:

From the main EViews menu bar select

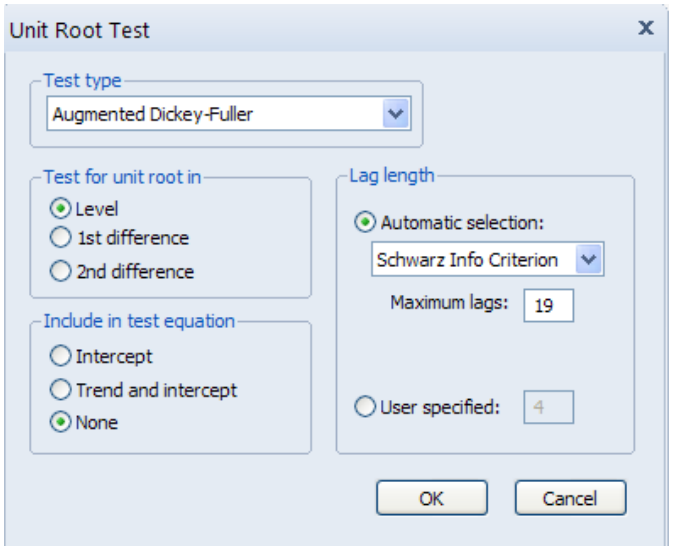
**Quick** **Series Statistics** **Unit Root Test**

Specify **copper** then click **OK**
In the ensuing dialog box, specify

| | |
|---|---|
| **Test Type:** | **Augmented Dickey-Fuller** |
| **Test for unit root in:** | **Level** |
| **Include in test equation:** | **None** |

This basic Dickey-Fuller test involves estimation of the regression

$$\Delta Y_t = \gamma\ Y_{t-1} + \varepsilon_t$$




The results window shows the regression output at the bottom. The *t*-statistic for γ is also shown at the top left of the window and the corresponding critical values are shown in the top right.

What is the test's null hypothesis? Can it be rejected for the exchange rate series? Is the series stationary?

(iii) Include a constant and a trend in the basic Dickey-Fuller test:

From the menu bar of the Dickey-Fuller results window, select

**View Unit Root Test**

The Unit Root Test specification dialog box should reappear. Specify

**Include in test equation: Trend and intercept**

Does the test indicate stationarity?

(iv) Carry out Augmented Dickey-Fuller tests, including lagged values of the dependent variable.

$$\Delta Y_t = a_0 + a_2 t + \gamma Y_{t-1} + \sum_{i=1}^{m} \beta_i \Delta Y_{t-i} + \varepsilon_t$$

If the extra terms are not significant, throw them out of the regression.

As in regression analysis, when you have reasonable *t*-values you are not finished; you must also check for autocorrelation in the residuals.

You can do this by selecting from the main menu bar:

**Quick Series Statistics Correlogram**

Then specify **resid**

Do the Augmented Dickey-Fuller tests indicate that the series is stationary?

(v) If the series is non-stationary, examine whether the first differences are stationary. Graph the differenced series, plot its correlogram and then perform unit root tests. (The difference operator is D(copper) in EViews.)

## (b) Arima Modelling

In the file ***session 4 workshop.wf1***, which you can download from the intranet, you have the data on the ***Gold*** prices from 1960 to 2018 (Dec).

The Box-Jenkins methodology first requires that we establish stationarity, before embarking on the three main steps: Identification, Estimation and Testing.

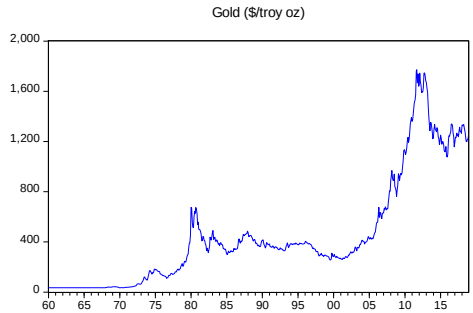
### (i) Establishing Stationarity

**Graph the Gold** time series by selecting **Quick** and then **Graph**. **Plot the autocorrelation function for the Gold** time series. The persistence of autocorrelations at long lags is typical of a non-stationary series. Another name for the autocorrelation function is the correlogram.

From the main EViews menu, select

**Quick Series Statistics Correlogram**

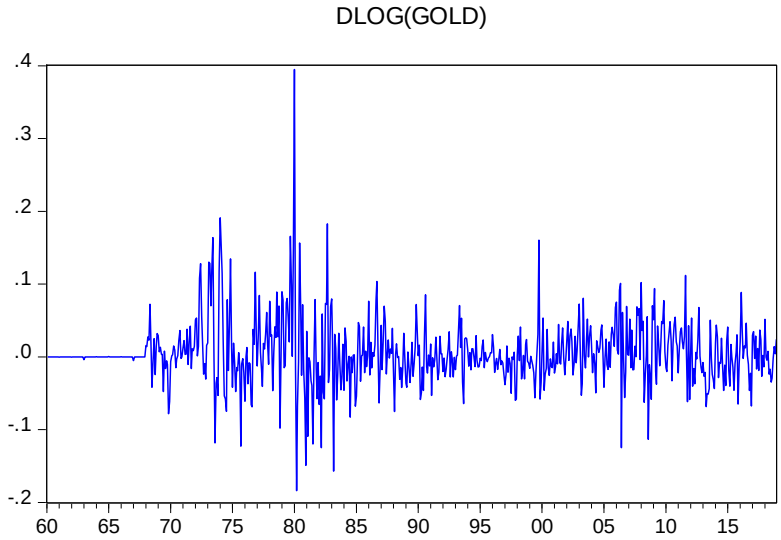
Specify the series as **Gold** and click **OK.**




Sample: 1960M01 2018M12
Included observations: 708

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.995 | 0.995 | 704.08 | 0.000 |
| | | 2 | 0.990 | -0.052 | 1401.6 | 0.000 |
| | | 3 | 0.985 | 0.023 | 2092.8 | 0.000 |
| | | 4 | 0.980 | 0.008 | 2778.1 | 0.000 |
| | | 5 | 0.975 | -0.010 | 3457.1 | 0.000 |
| | | 6 | 0.969 | -0.071 | 4129.2 | 0.000 |
| | | 7 | 0.963 | -0.042 | 4793.6 | 0.000 |
| | | 8 | 0.956 | -0.028 | 5450.1 | 0.000 |
| | | 9 | 0.950 | -0.012 | 6098.5 | 0.000 |
| | | 10 | 0.943 | 0.008 | 6739.2 | 0.000 |
| | | 11 | 0.937 | -0.005 | 7372.1 | 0.000 |
| | | 12 | 0.930 | -0.043 | 7996.5 | 0.000 |
| | | 13 | 0.923 | 0.022 | 8612.9 | 0.000 |
| | | 14 | 0.916 | -0.013 | 9221.0 | 0.000 |
| | | 15 | 0.909 | -0.005 | 9820.9 | 0.000 |
| | | 16 | 0.902 | -0.043 | 10412. | 0.000 |
| | | 17 | 0.895 | -0.009 | 10995. | 0.000 |
| | | 18 | 0.888 | 0.017 | 11569. | 0.000 |
| | | 19 | 0.880 | -0.023 | 12134. | 0.000 |
| | | 20 | 0.873 | 0.018 | 12692. | 0.000 |
| | | 21 | 0.866 | -0.028 | 13240. | 0.000 |
| | | 22 | 0.859 | 0.027 | 13780. | 0.000 |
| | | 23 | 0.851 | -0.010 | 14312. | 0.000 |
| | | 24 | 0.844 | 0.015 | 14836. | 0.000 |
| | | 25 | 0.837 | 0.006 | 15353. | 0.000 |
| | | 26 | 0.830 | -0.040 | 15860. | 0.000 |
| | | 27 | 0.823 | -0.003 | 16360. | 0.000 |
| | | 28 | 0.815 | -0.064 | 16851. | 0.000 |
| | | 29 | 0.807 | 0.005 | 17333. | 0.000 |
| | | 30 | 0.800 | 0.010 | 17807. | 0.000 |

**Graph the log returns of the series, DLOG(gold).** The series looks stationary. **Plot the autocorrelation function of DLOG(gold).** The autocorrelations are not very persistent which suggests stationarity so let's build a model for DLOG(gold). (If Dickey-Fuller tests were performed, they would support stationarity in the mean.)



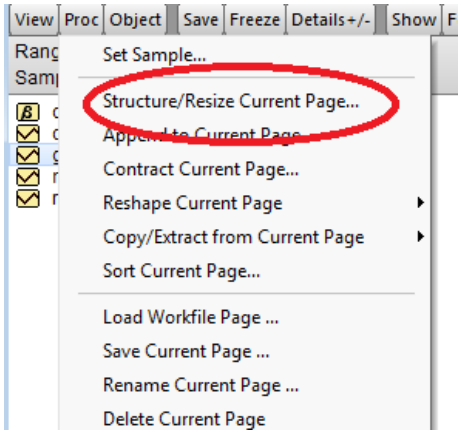

Sample: 1960M01 2018M12
Included observations: 707

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.262 | 0.262 | 48.914 | 0.000 |
| | | 2 | -0.039 | -0.116 | 49.995 | 0.000 |
| | | 3 | 0.016 | 0.063 | 50.175 | 0.000 |
| | | 4 | 0.062 | 0.039 | 52.931 | 0.000 |
| | | 5 | 0.001 | -0.025 | 52.932 | 0.000 |
| | | 6 | -0.047 | -0.035 | 54.539 | 0.000 |
| | | 7 | 0.099 | 0.129 | 61.525 | 0.000 |
| | | 8 | 0.141 | 0.073 | 75.682 | 0.000 |
| | | 9 | 0.101 | 0.067 | 82.974 | 0.000 |
| | | 10 | 0.065 | 0.042 | 85.993 | 0.000 |
| | | 11 | 0.150 | 0.133 | 102.16 | 0.000 |
| | | 12 | 0.119 | 0.041 | 112.35 | 0.000 |
| | | 13 | -0.041 | -0.065 | 113.56 | 0.000 |
| | | 14 | -0.096 | -0.068 | 120.22 | 0.000 |
| | | 15 | 0.019 | 0.032 | 120.48 | 0.000 |
| | | 16 | 0.067 | 0.016 | 123.78 | 0.000 |
| | | 17 | -0.003 | -0.028 | 123.79 | 0.000 |
| | | 18 | -0.052 | -0.071 | 125.79 | 0.000 |
| | | 19 | -0.021 | -0.054 | 126.11 | 0.000 |
| | | 20 | 0.066 | 0.043 | 129.29 | 0.000 |
| | | 21 | -0.037 | -0.073 | 130.28 | 0.000 |
| | | 22 | 0.012 | 0.061 | 130.38 | 0.000 |
| | | 23 | 0.038 | -0.006 | 131.41 | 0.000 |
| | | 24 | -0.008 | -0.018 | 131.46 | 0.000 |

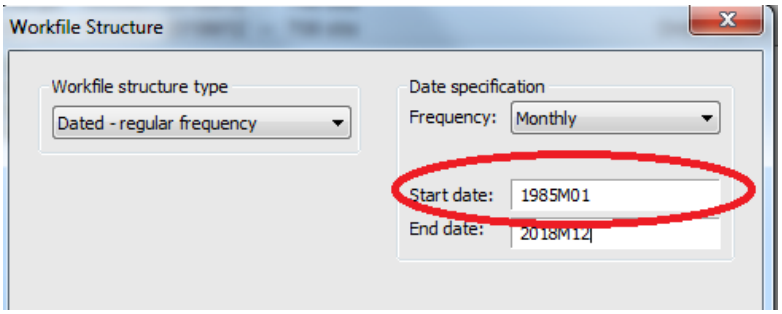
**You may want to remove the initial data, as it is not representative**.

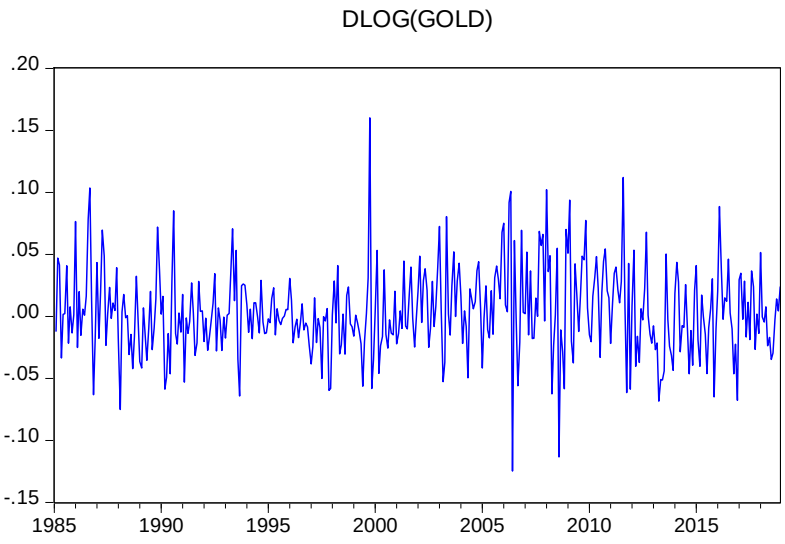
From the Workfile Window menu, select

**Proc Structure/Resize Current Page**




Set the new workfile structure type to start from 1985M01



**(ii) Identification**

Plot the autocorrelation function (ACF) and partial autocorrelation function (PACF) for the series DLOG(gold).

Both of these can be found by selecting from the main EViews menu

**Quick Series Statistics Correlogram**

Sample: 1985M01 2018M12
Included observations: 407

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.136 | 0.136 | 7.6014 | 0.006 |
| | | 2 | -0.085 | -0.106 | 10.580 | 0.005 |
| | | 3 | -0.000 | 0.028 | 10.580 | 0.014 |
| | | 4 | 0.047 | 0.035 | 11.506 | 0.021 |
| | | 5 | 0.018 | 0.009 | 11.647 | 0.040 |
| | | 6 | 0.005 | 0.009 | 11.657 | 0.070 |
| | | 7 | 0.072 | 0.074 | 13.829 | 0.054 |
| | | 8 | 0.024 | 0.001 | 14.060 | 0.080 |
| | | 9 | -0.013 | -0.004 | 14.130 | 0.118 |
| | | 10 | 0.009 | 0.013 | 14.163 | 0.166 |
| | | 11 | 0.154 | 0.148 | 24.152 | 0.012 |
| | | 12 | 0.030 | -0.017 | 24.534 | 0.017 |
| | | 13 | -0.001 | 0.027 | 24.534 | 0.027 |
| | | 14 | 0.035 | 0.027 | 25.056 | 0.034 |
| | | 15 | 0.045 | 0.027 | 25.923 | 0.039 |
| | | 16 | 0.042 | 0.034 | 26.664 | 0.045 |
| | | 17 | 0.019 | 0.015 | 26.820 | 0.061 |
| | | 18 | 0.057 | 0.038 | 28.209 | 0.059 |
| | | 19 | 0.021 | 0.002 | 28.390 | 0.076 |
| | | 20 | 0.081 | 0.089 | 31.181 | 0.053 |
| | | 21 | 0.027 | -0.002 | 31.496 | 0.066 |
| | | 22 | 0.036 | 0.020 | 32.060 | 0.076 |
| | | 23 | 0.037 | 0.025 | 32.656 | 0.087 |
| | | 24 | 0.047 | 0.037 | 33.608 | 0.092 |
| | | 25 | 0.009 | -0.016 | 33.647 | 0.116 |
| | | 26 | 0.062 | 0.064 | 35.321 | 0.105 |
| | | 27 | -0.010 | -0.056 | 35.367 | 0.130 |
| | | 28 | 0.001 | 0.013 | 35.368 | 0.159 |
| | | 29 | -0.001 | -0.030 | 35.368 | 0.193 |
| | | 30 | -0.024 | -0.031 | 35.615 | 0.221 |
| | | 31 | 0.052 | 0.025 | 36.817 | 0.218 |
| | | 32 | -0.005 | -0.033 | 36.829 | 0.255 |
| | | 33 | -0.043 | -0.052 | 37.634 | 0.265 |
| | | 34 | -0.047 | -0.056 | 38.641 | 0.268 |
| | | 35 | 0.051 | 0.038 | 39.790 | 0.265 |
| | | 36 | -0.032 | -0.073 | 40.258 | 0.287 |

The PACF and ACF function have two significant spikes and then cut-off (even though the ACF seems to fading away). However, we still have some spikes appearing later on, and also at lags 7, 11 and 20.

A number of ARMA processes could result in this pattern. Verify that reasonable candidate models are: ARMA(2,0), ARMA(0,2), ARMA(1,0), ARMA(0,1), and ARMA(1,1).

**(iii) Estimation**

Models can be estimated by selecting from the main EViews menu

**Quick Estimate Equation**

and in the Equation Specification dialog box, we might choose to fit an ARMA(0,1)

**DLOG(gold) c MA(1)**

**(iv) Testing**

Both of the coefficients are significant, so the model looks quite reasonable. However, as in regression analysis, we must consider the residuals.

The quickest way to graph the residuals is to select from the menu bar of the equation window

**View Actual, Fitted, Residual Actual, Fitted, Residual Graph**

Do the **residuals** look random?

We should check for **autocorrelation in the residuals**. From the menu bar of the equation window, select

**View Residual Diagnostics Correlogram**

Does the correlogram suggest random residuals?

The **Box-Pierce Q-statistic** is shown to the right of the ACF and PACF plots. This statistic considers autocorrelation in several lags together. The p-value of the Q-statistic is given for convenience. When testing at 5% significance, a p-value less than 5% indicates significance. Are any of the Q-statistics significant?

Sample: 1985M01 2018M12
Included observations: 407
Q-statistic probabilities adjusted for 1 ARMA term

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | -0.014 | -0.014 | 0.0774 | |
| | | 2 | -0.083 | -0.084 | 2.9417 | 0.086 |
| | | 3 | 0.006 | 0.004 | 2.9572 | 0.228 |
| | | 4 | 0.044 | 0.037 | 3.7588 | 0.289 |
| | | 5 | 0.012 | 0.015 | 3.8229 | 0.431 |
| | | 6 | -0.009 | -0.002 | 3.8558 | 0.570 |
| | | 7 | 0.071 | 0.073 | 5.9516 | 0.429 |
| | | 8 | 0.014 | 0.013 | 6.0305 | 0.536 |
| | | 9 | -0.013 | -0.002 | 6.0984 | 0.636 |
| | | 10 | -0.014 | -0.013 | 6.1854 | 0.721 |
| | | 11 | 0.155 | 0.150 | 16.262 | 0.092 |
| | | 12 | 0.006 | 0.005 | 16.277 | 0.131 |
| | | 13 | -0.007 | 0.020 | 16.296 | 0.178 |
| | | 14 | 0.030 | 0.028 | 16.689 | 0.214 |
| | | 15 | 0.034 | 0.025 | 17.181 | 0.247 |
| | | 16 | 0.035 | 0.037 | 17.715 | 0.278 |
| | | 17 | 0.004 | 0.014 | 17.723 | 0.340 |
| | | 18 | 0.056 | 0.042 | 19.070 | 0.325 |
| | | 19 | -0.001 | -0.006 | 19.071 | 0.387 |
| | | 20 | 0.079 | 0.087 | 21.741 | 0.297 |
| | | 21 | 0.009 | 0.011 | 21.778 | 0.353 |
| | | 22 | 0.030 | 0.016 | 22.175 | 0.389 |
| | | 23 | 0.025 | 0.021 | 22.446 | 0.434 |
| | | 24 | 0.044 | 0.045 | 23.295 | 0.444 |
| | | 25 | -0.009 | -0.021 | 23.332 | 0.500 |
| | | 26 | 0.067 | 0.069 | 25.267 | 0.447 |
| | | 27 | -0.022 | -0.047 | 25.478 | 0.492 |
| | | 28 | 0.003 | 0.007 | 25.483 | 0.547 |
| | | 29 | 0.004 | -0.022 | 25.491 | 0.601 |
| | | 30 | -0.034 | -0.040 | 25.989 | 0.626 |
| | | 31 | 0.059 | 0.024 | 27.526 | 0.596 |
| | | 32 | -0.009 | -0.021 | 27.565 | 0.644 |
| | | 33 | -0.032 | -0.046 | 28.030 | 0.668 |
| | | 34 | -0.052 | -0.071 | 29.253 | 0.654 |
| | | 35 | 0.065 | 0.038 | 31.144 | 0.608 |
| | | 36 | -0.033 | -0.054 | 31.638 | 0.631 |

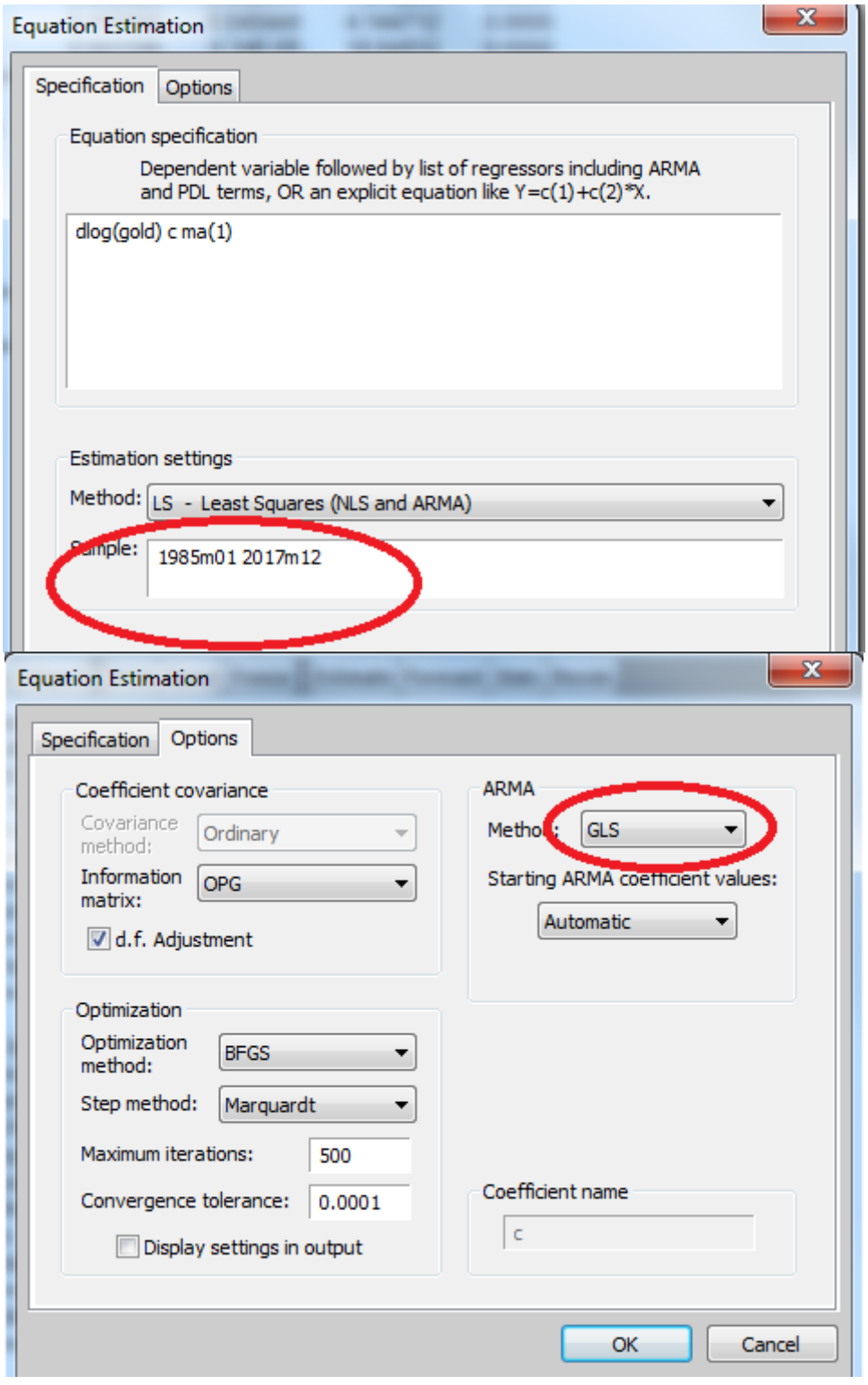
Take note of the **RSS**, **SBC, AIC,** and **HQ**; these indicate the quality of the fit of the model: the lower the better. SBC, AIC and HQ can be negative.

Try to find an alternative model which has better or comparable fit, and which also has no residual autocorrelation. The quickest way to estimate another model is to select **Estimate** from the menu bar of the equation window; the Equation Specification dialog box should reappear.

ARMA(2, 0): **DLOG(gold) c AR(1) AR(2)**
ARMA(0, 2): **DLOG(gold) c MA(1) MA(2)**
ARMA(1, 0): **DLOG(gold) c AR(1)**
ARMA(1, 1): **DLOG(gold) c AR(1) MA(1)**

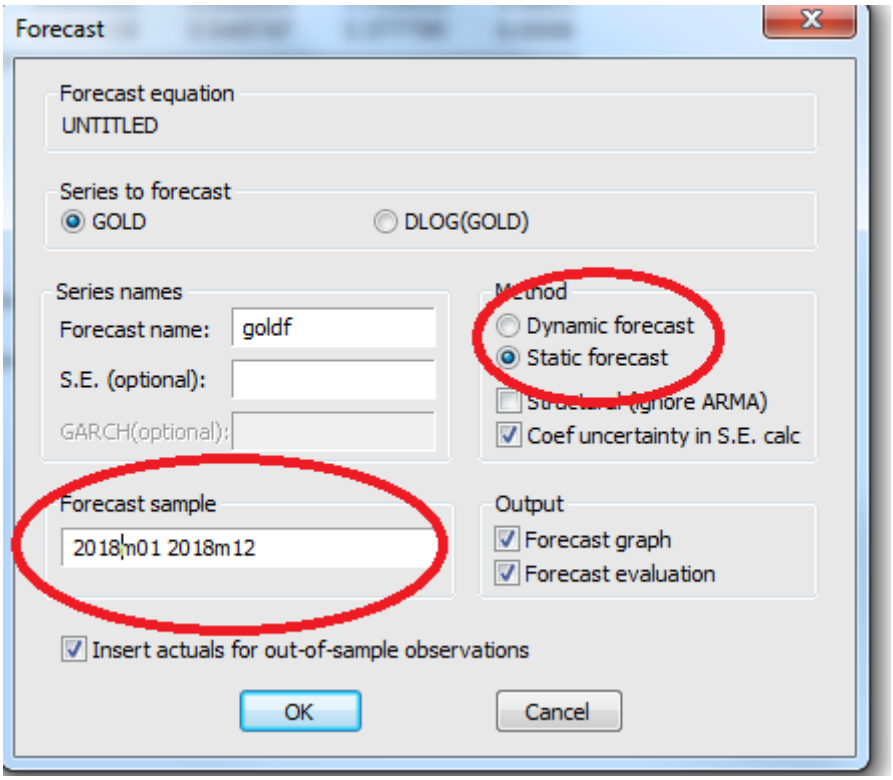
**(v) Forecasting**

Estimate an ARMA(0,1) model for the DLOG(gold) series using all the data **except the last 12 observations**. You can specify the shortened sample at the bottom of the equation specification dialog box.




Click OK. Check that the model has reasonable diagnostics.

Click the **Forecast** button which can be found in the menu at the top of the current equation window.




In the top right of the dialog box, change the forecasting **Method** to **Static** and click OK.

Now try a **Dynamic** forecast for gold.

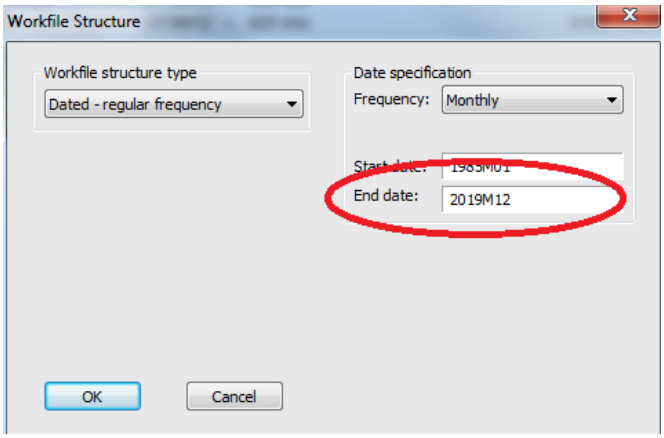
Both Dynamic and Static forecasting produce 12 forecasts but they are not the same. Dynamic produces a 1-step-ahead forecast, a 2-step-ahead forecast, a 3-step-ahead forecast, ……, and a 12-step-ahead forecast. By contrast, Static forecasting produces twelve 1-step-ahead forecasts.

Produce Static forecasts for Gold and then plot these forecasts, goldF, and gold on the same graph by selecting **Quick Graph**. How good is the fit?

Produce Static forecasts for DLOG(gold) by specifying this at the top left of the forecast dialog box. Plot the forecasts and DLOG(gold) on the same graph. How good is the fit?

***Do you expect gold prices to go up or down in the next 12 months*?**
**How about risk? What is the prediction interval?**

Click: ***workfile window … Procs … Structure/Resize Current Page***
Date Specification. **End Date: 2019:M12**




Once again, estimate an ARMA(0,1) model for the DLOG(gold) series, but now using all the data (**2018M12**).

Click the **Forecast** button which can be found in the menu at the top of the current equation window. [FORECAST observations **2019m01 to 2019m12**]. In the top right of the dialog box, try a **Dynamic** forecasting **Method** and click OK.

Why can't you use the **Static** Forecasting method?