# DSC5211C QUANTITATIVE RISK MANAGEMENT

## SESSION 4

**Fernando Oliveira**
bizfmdo@nus.edu.sg

## ARIMA Modelling

# Objectives

- Box-Jenkins Univariate Forecasting
    - Introduction to ARIMA Models
    - Model Identification
    - Model Estimation

Readings: Enders, Ch. 2.6-2.10
Franses, Ch. 3
Hanke, Ch. 9.

# Box-Jenkins Univariate Forecasting

- Box-Jenkins is a univariate forecasting approach. It involves careful examination of series in order to identify underlying ***data-generating process***.
- Useful to restrict search in models of the class of **A**uto**R**egressive **I**ntegrated **M**oving **A**verage: these models are linear functions of past observations: $Y_t = 0.5\, Y_{t-1} + 0.3\, Y_{t-2} + e_t$
- Choice of best model can be made systematically by following sequence of:

MODEL IDENTIFICATION

↓

MODEL ESTIMATION

↓

DIAGNOSTIC EVALUATION

↓

Box and Jenkins (1970)

# Autoregressive Models

- If $Y_t$ follows an *autoregressive* process of order $p$, we write $Y_t$ ~ AR($p$), or $Y_t$ ~ ARMA($p$,0)

  $Y_t$ depends linearly on $p$ past values:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ..... + \phi_p Y_{t-p} + e_t$$

where $e_t$ has zero mean, fixed variance, no autocorrelation (as in regression)

# Autoregressive Models: AR(1)

- An AR(1) process:

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

- Unemployment, Prices, GDP, etc.
- AR model of sufficiently high order can generally be found as a close approximation to any process met in business and economics
- ***But if a large number of parameters are required for good fit, forecasts can be poor.***
- This motivates working with a broader class of models.

# Moving Average Models

- If $Y_t$ follows a moving average process of order $q$, we write

  $Y_t \sim$ MA($q$), or $Y_t \sim$ ARMA(0,$q$)

  $Y_t$ depends linearly on $q$ past values of a stochastic term:

$$Y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots.. + \theta_q e_{t-q}$$

  $e_t$ has zero mean, fixed variance, no autocorrelation

- Helpful not to view $e_t$ as ‘error’ but instead as a pure stochastic term affecting $Y_t$

# Moving Average Models: MA(1)

- An MA(1) process:

$$Y_t = e_t + \theta_1 e_{t-1}$$

- Small commodity market receives news about crops. News will have immediate effect and discounted effect as market assimilates importance.

- Don't confuse this MA with the smoothing MA

# ARMA Models

- An obvious generalisation of AR and MA models, that includes them as special cases, is the *mixed* model in which $Y_t$ is generated by

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots.. + \phi_p Y_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots.. + \theta_q e_{t-q}$$

  This is an ARMA($p$,$q$) model

- Experience suggests that a mixed model ARMA($p$,$q$) may achieve as good a fit as an AR($p'$) model but using fewer parameters,

  i.e. $p+q<p'$

- Since amount of data is limited, preference is to fit a model involving as few parameters as possible

  This is known as the ***principle of parsimony***

# ARIMA Models

- Univariate methods rely on information in past series for predicting future.

- The assumption is that there is some regularity in data generating process.

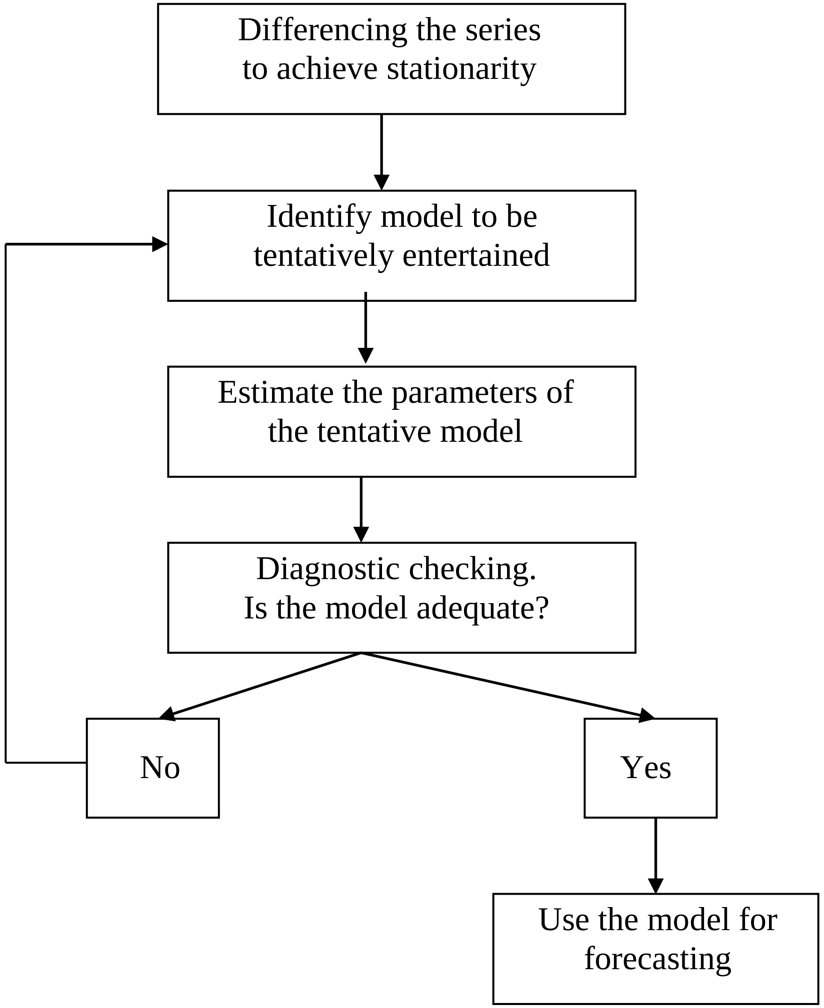
- Assuming stability over time of autocorrelations allows their exploitation in forecasting

- Use of autocorrelation structure relies on series being stationary:
    - constant mean
    - constant variance
    - autocorrelation depends only on lag

# ARIMA Models

- Use of autocorrelation structure relies on series being stationary:
- Stationarity is achieved by differencing. This is consistent with current view that trends in financial and economic series are stochastic.
- If variable must be differentiated $d$ times in order to achieve stationarity, it is *integrated* of order $d$
- If it then follows an ARMA($p$,$q$) process, we say that the variable is an ARIMA($p$,$d$,$q$) process: **A**uto**R**egressive **I**ntegrated **M**oving **A**verage

# Box-Jenkins Methodology



# Model Identification (*p*,*d*,*q*)

- Box-Jenkins imposes stationarity by differencing.

  If differenced $d$ times, model will be ARIMA($p$,$d$,$q$)

- After differencing, look at ACF and PACF of new stationary series.

  Different ARIMA models have different ACF and PACF patterns.

- *Autocorrelation function* (ACF): measures correlation between $Y_t$ and $Y_{t-k}$

## ACF & PACF for AR Models

- To get a feel for ACF & PACF, consider

$$Y_t = 0.8Y_{t-1} + \varepsilon_t \qquad (1)$$

- Using this:

$$Y_{t-1} = 0.8Y_{t-2} + \varepsilon_{t-1} \qquad (2)$$

Substituting (2) into (1):

$$Y_t = 0.64Y_{t-2} + 0.8\varepsilon_{t-1} + \varepsilon_t \quad (3)$$

# ACF & PACF for AR Models

Using (1):

$$corr(Y_t, Y_{t-1}) = corr(0.8Y_{t-1} + \varepsilon_t, Y_{t-1})$$

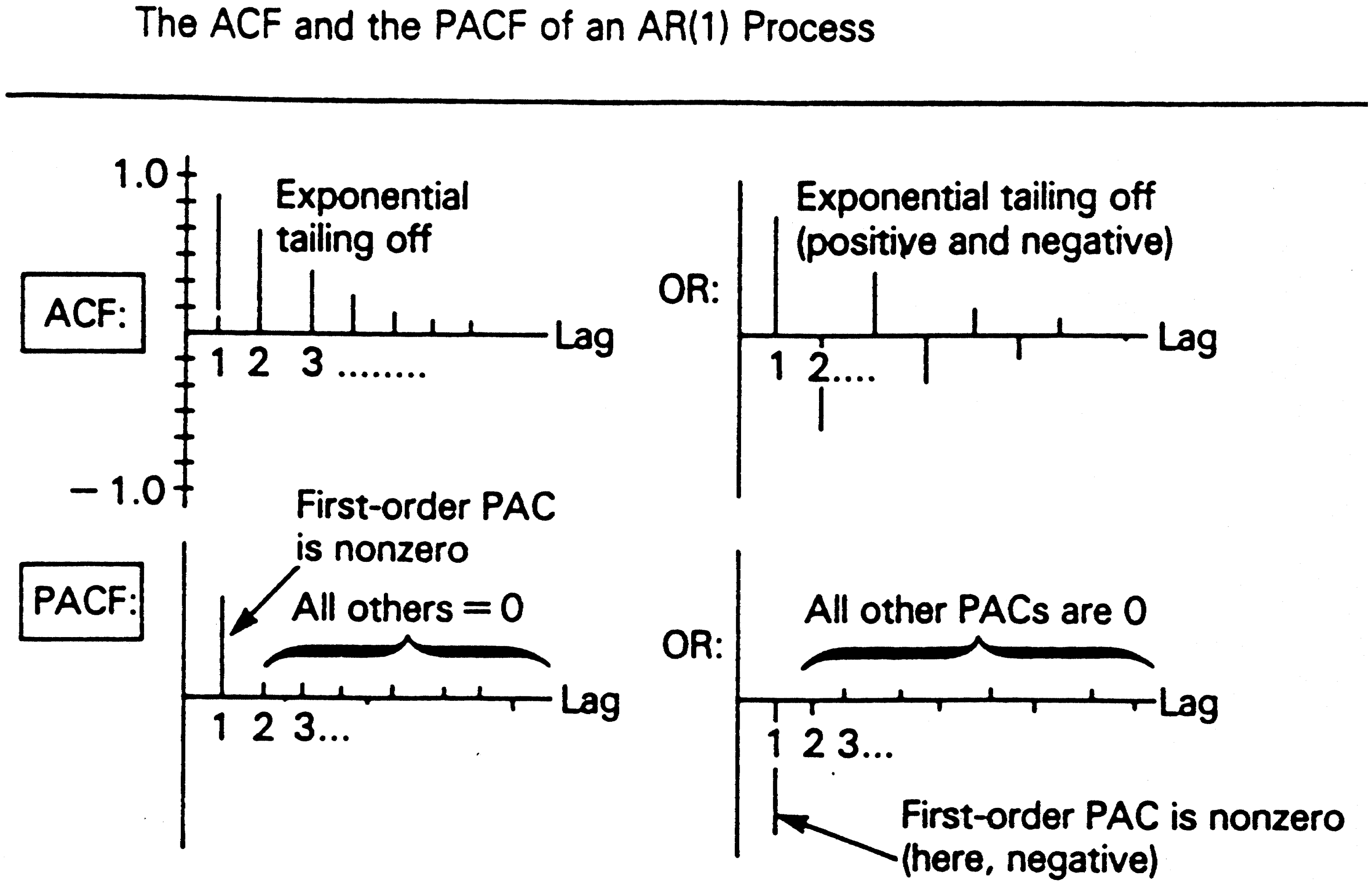
Using (3):

$$corr(Y_t, Y_{t-2}) = corr(0.64Y_{t-2} + 0.8\varepsilon_{t-1} + \varepsilon_t, Y_{t-2})$$

Indicates that autocorrelations are significant for first few orders and die away exponentially

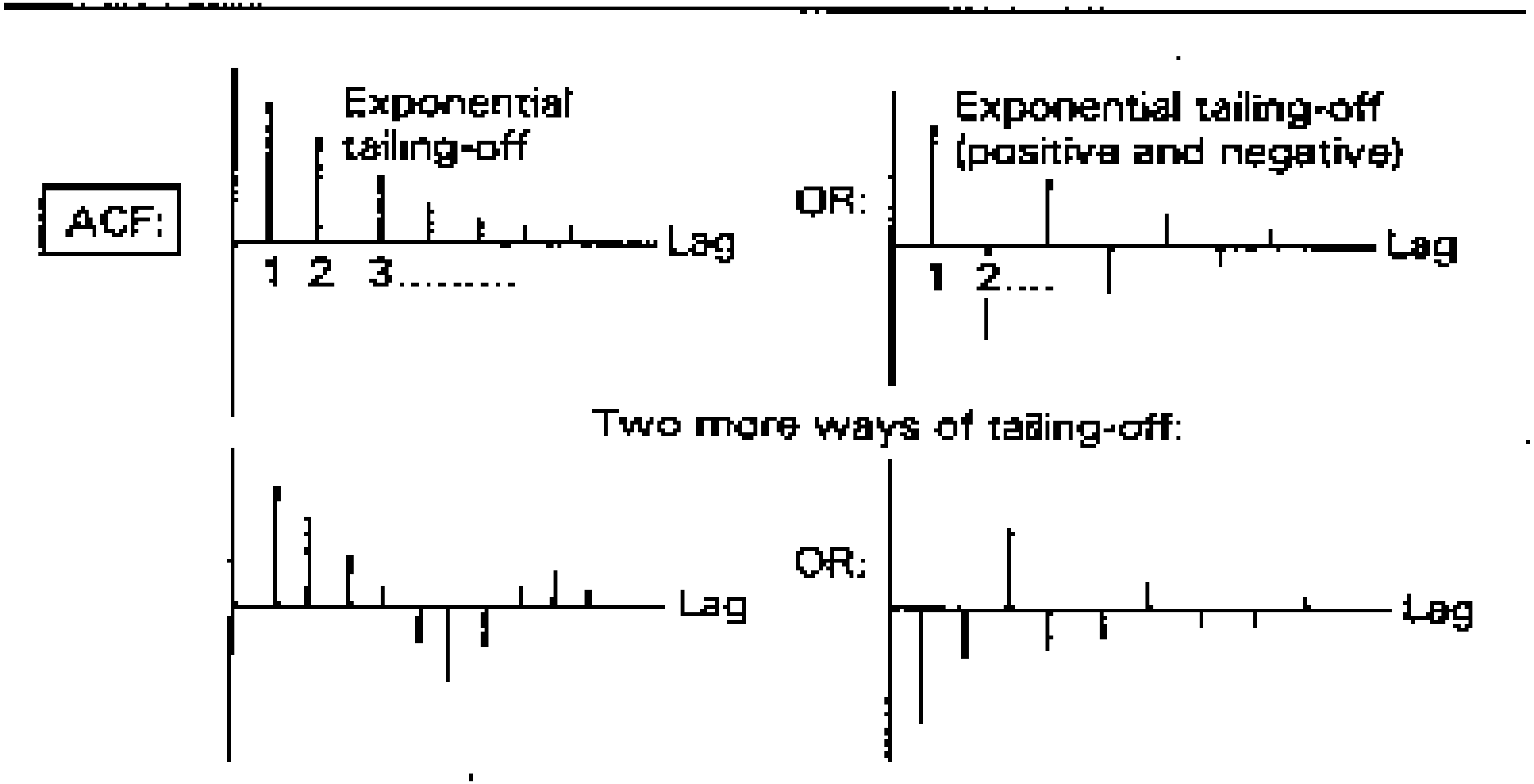
- To get a feel for PACF, ask how significant is $Y_{t-k}$ in a regression of $Y_t$ on $Y_{t-1}$, $Y_{t-2}$, $Y_{t-3}$,….., $Y_{t-k}$

# ACF & PACF for AR(1) Model

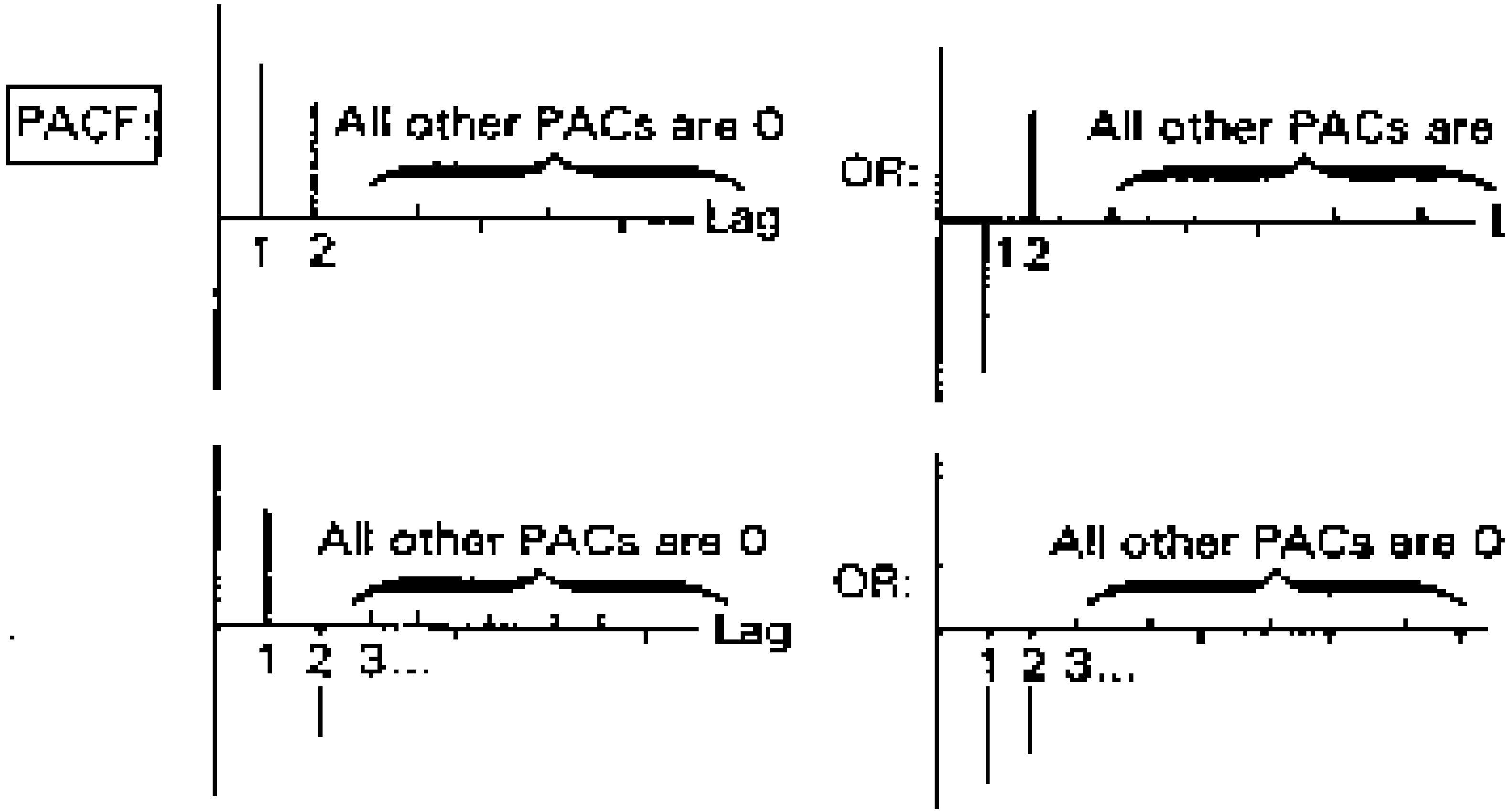


# ACF & PACF for AR(2) Model

The ACF and the PACF of an AR(2) Process



# ACF & PACF for AR(2) Model




In each case, the first two PAC's are nonzero, then the PACF cuts off to zero

# ACF & PACF for MA Models

- To get a feel for ACF & PACF, consider

$$Y_t = 0.8\varepsilon_{t-1} + \varepsilon_t \qquad (1)$$

- From (1):

$$Y_{t-1} = 0.8\varepsilon_{t-2} + \varepsilon_{t-1} \qquad (2)$$

$$Y_{t-2} = 0.8\varepsilon_{t-3} + \varepsilon_{t-2} \qquad (3)$$

- Using (1) and (2):

$$corr(Y_t, Y_{t-1}) = corr(0.8\varepsilon_{t-1} + \varepsilon_t, 0.8\varepsilon_{t-2} + \varepsilon_{t-1})$$

- Using (1) and (3): Indicates that autocorrelation is significant for first order but is zero thereafter

# ACF & PACF for MA Models

- To calculate PACF, express RHS in terms of $Y_{t-1}$, $Y_{t-2}$, $Y_{t-3}$, ….., $Y_{t-k}$

From (2): $\varepsilon_{t-1} = Y_{t-1} - 0.8\varepsilon_{t-2}$

So from (1): $Y_t = 0.8Y_{t-1} - 0.64\varepsilon_{t-2} + \varepsilon_t$ (4)

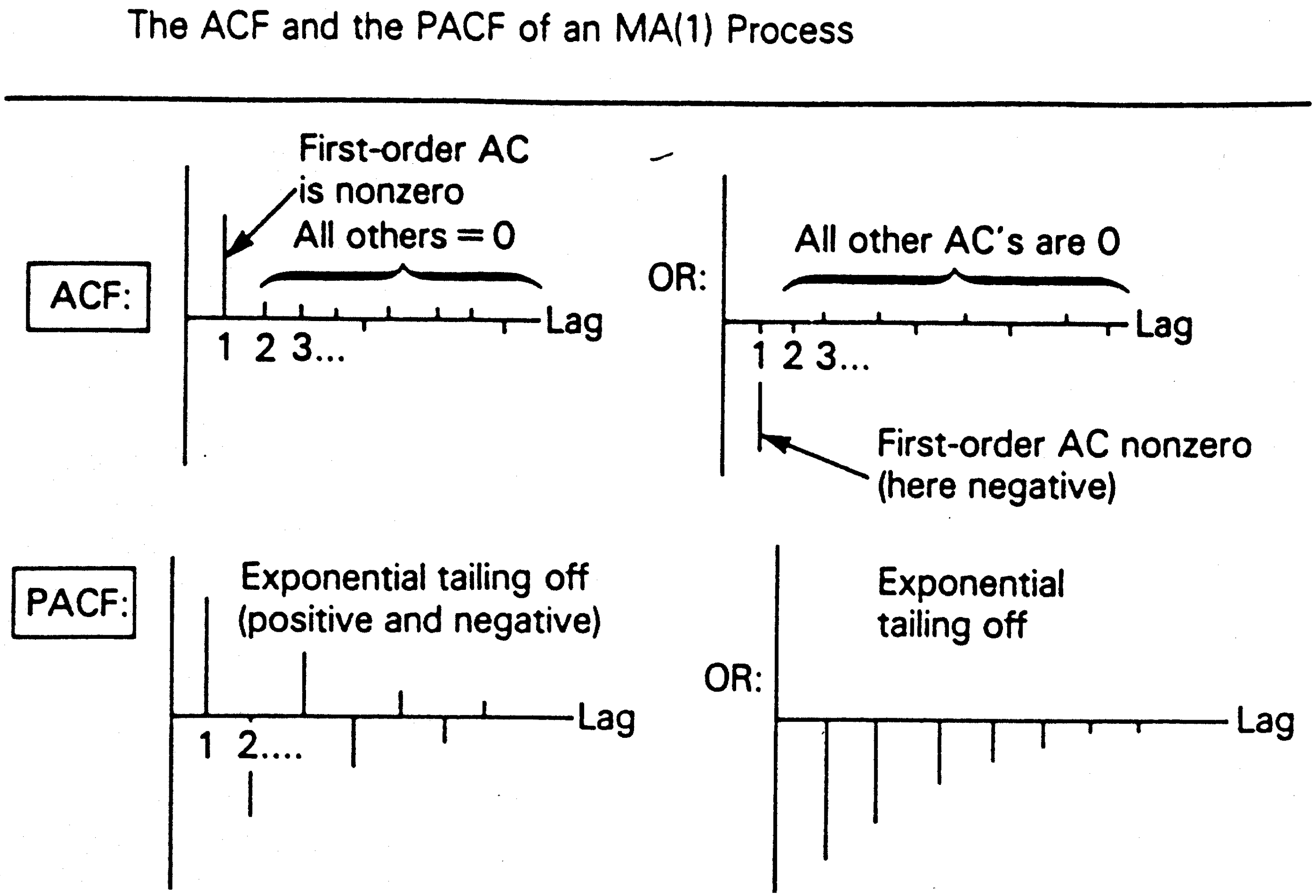
From (3): $\varepsilon_{t-2} = Y_{t-2} - 0.8\varepsilon_{t-3}$

So from (4): $Y_t = 0.8Y_{t-1} - 0.64Y_{t-2} + 0.512\varepsilon_{t-3} + \varepsilon_t$

Thus $Y_t = 0.8Y_{t-1} - 0.64Y_{t-2} + 0.512Y_{t-3} + ....... + \varepsilon_t$

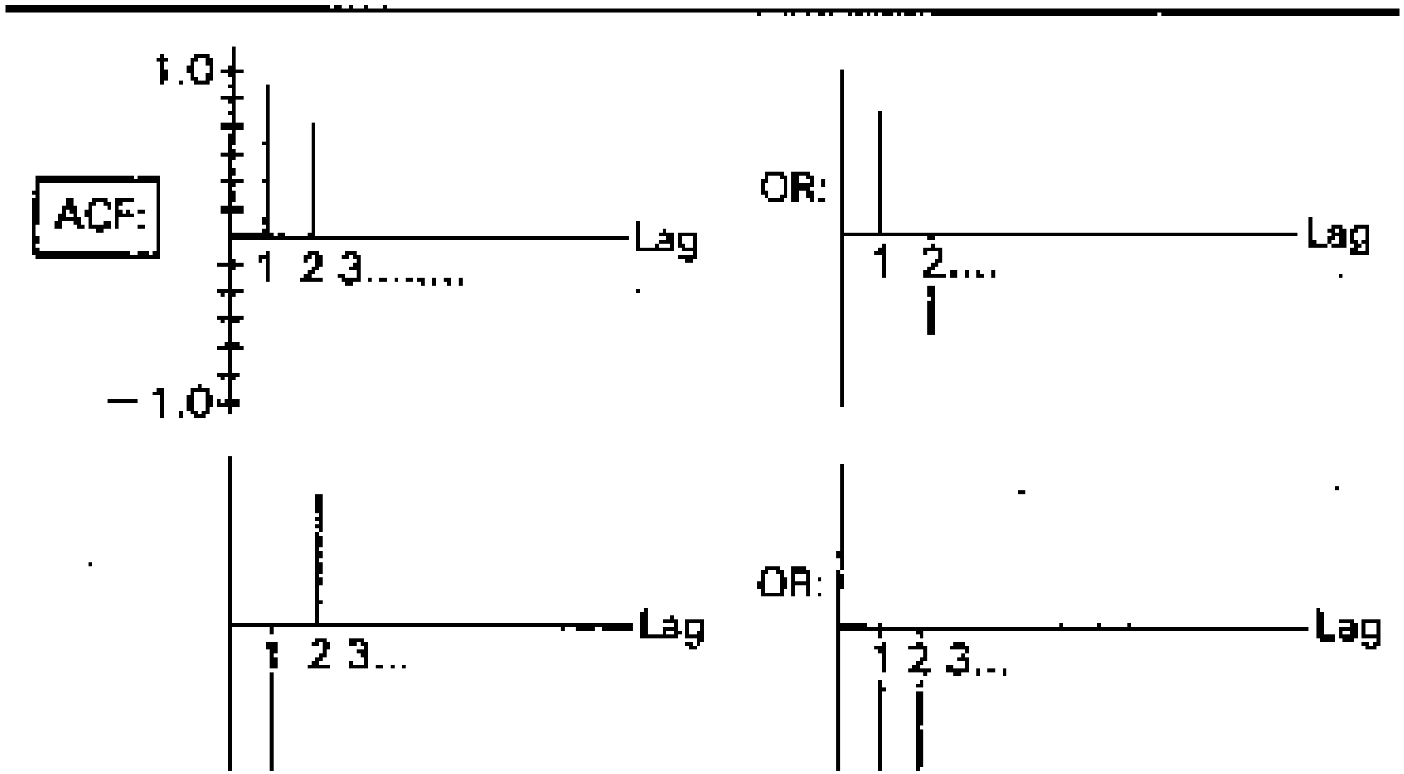
Indicates PACF significant for first few orders and then dies away exponentially

# ACF & PACF for MA(1)



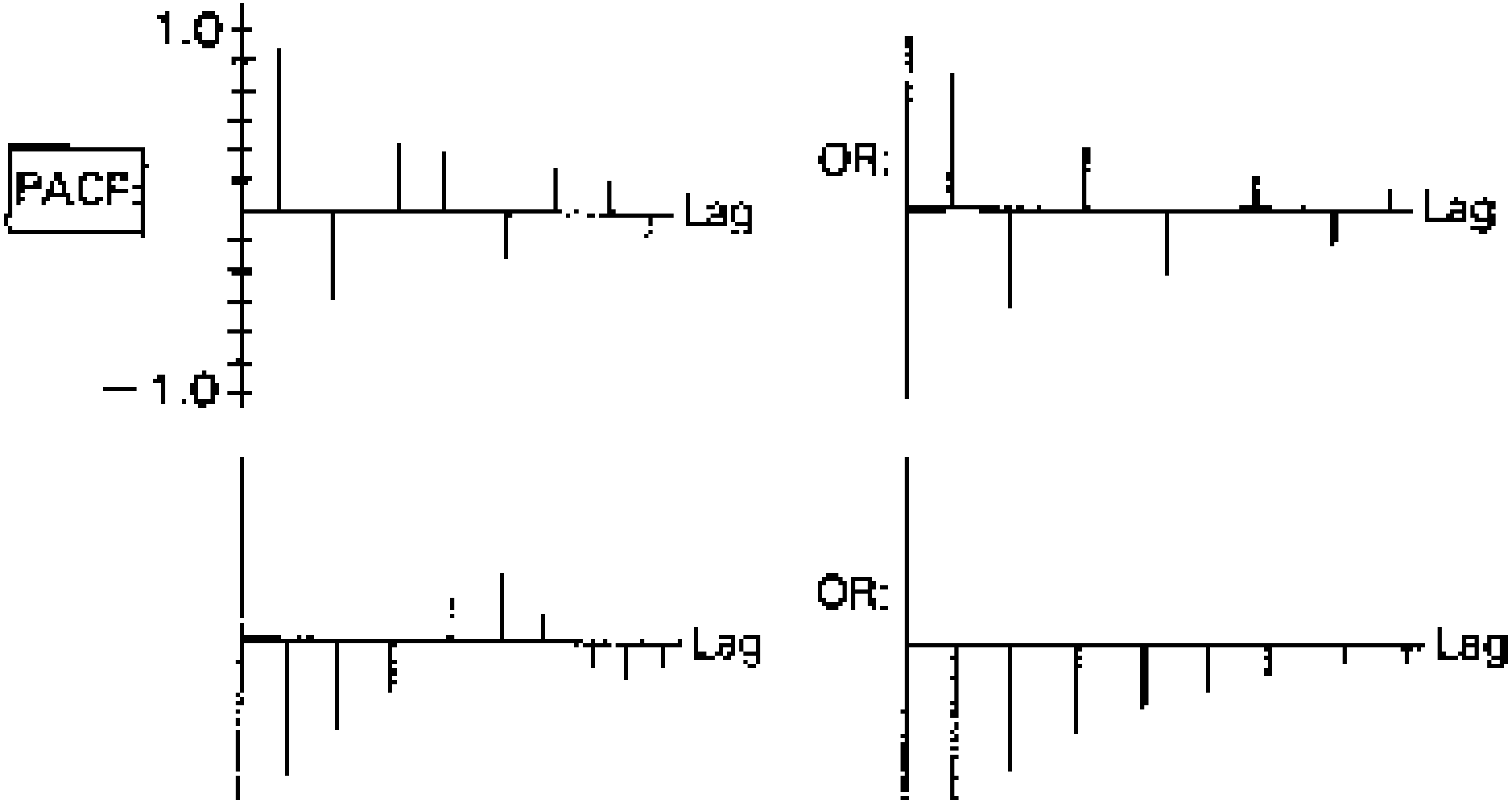
# ACF & PACF for MA(2)

The ACF and the PACF of an MA(2) Process




The first two autocorrelations (or at least the second one) are nonzero, either positive or negative, and all the rest are zero

# ACF & PACF for MA(2)

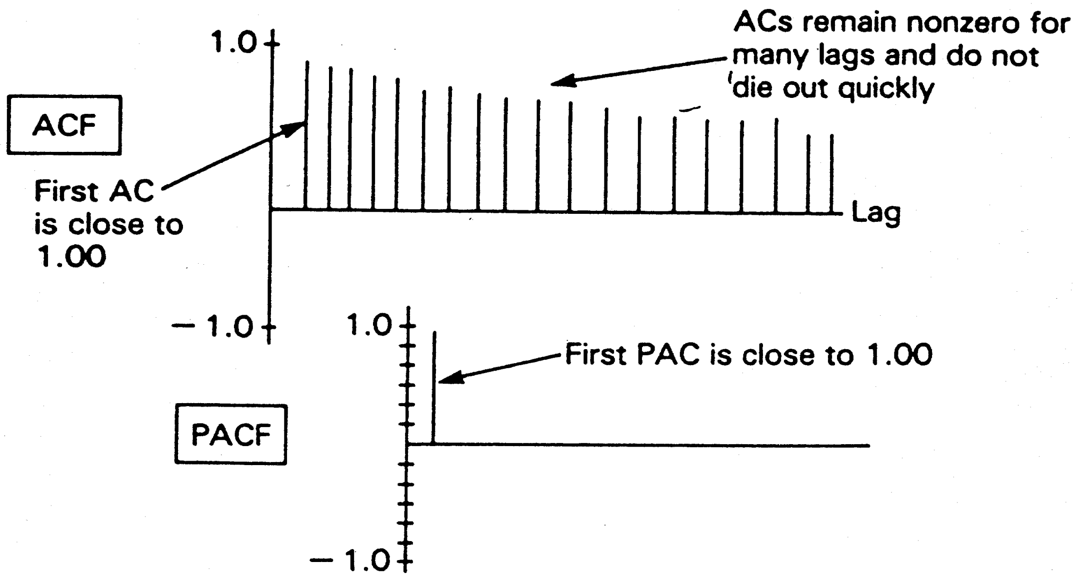



All partial autocorrelations tail off in some fashion, as shown

# ACF & PACF for Non-Stationary Models

- Differencing should be performed to make series stationary before one looks at ACF and PACF

- However, non-stationarity should be apparent from ACF and PACF



# Summary: ARIMA Identification

| **Type of model** | **ACF** | **PACF** |
|---|---|---|
| AR($p$) | Tails off | $p$ spikes, then cuts off to zero |
| MA($q$) | $q$ spikes, then cuts to zero | Tails off |
| Mixed ARMA | Tails off | Tails off |
| Nonstationary | Persistently nonzero | Large spike, close to 1.00, at lag 1 |

# Model Estimation

- For purely AR processes

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

parameters can be estimated by a regression package or a dedicated ARIMA package

- When the model contains MA parameters,

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

there is a need for a dedicated ARIMA package

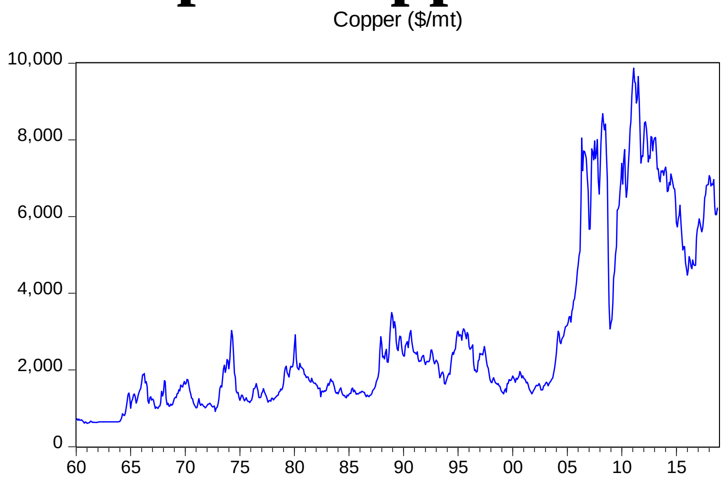
# Model Estimation

- *Maximum likelihood* - this procedure selects the parameters so as to maximise the likelihood of drawing the observed sample

- Amounts to minimisation of complicated nonlinear function of parameters.

- Iterative numerical procedure is used to search for optimal parameters.

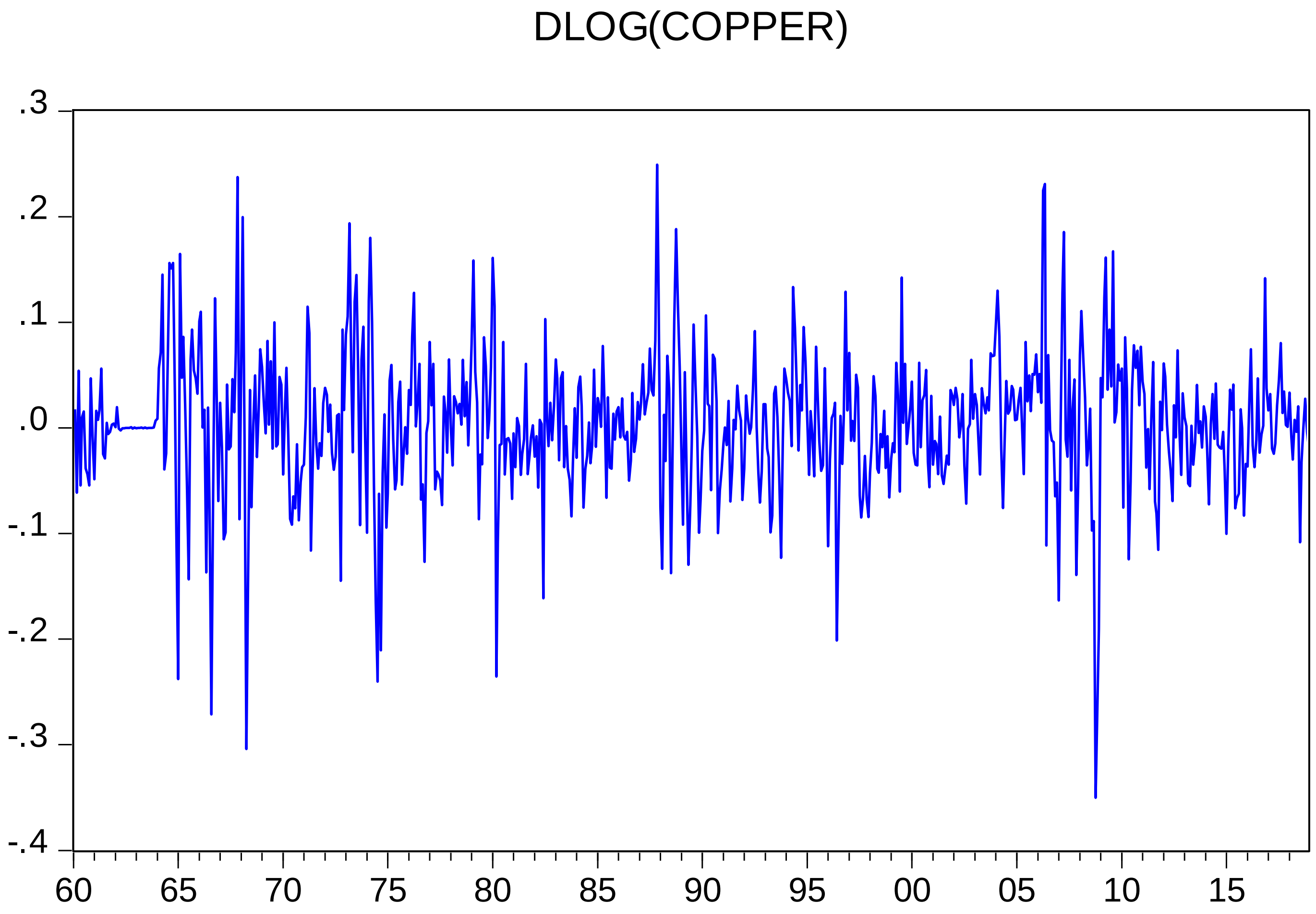
# Example: Copper Prices




Sample: 1960M01 2018M12
Included observations: 708

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.992 | 0.992 | 699.84 | 0.000 |
| | | 2 | 0.980 | -0.257 | 1384.0 | 0.000 |
| | | 3 | 0.968 | 0.008 | 2051.6 | 0.000 |
| | | 4 | 0.955 | 0.030 | 2703.3 | 0.000 |
| | | 5 | 0.944 | 0.004 | 3339.9 | 0.000 |
| | | 6 | 0.931 | -0.037 | 3961.0 | 0.000 |
| | | 7 | 0.919 | 0.027 | 4567.3 | 0.000 |
| | | 8 | 0.909 | 0.068 | 5160.5 | 0.000 |
| | | 9 | 0.901 | 0.111 | 5743.7 | 0.000 |
| | | 10 | 0.893 | -0.004 | 6318.5 | 0.000 |
| | | 11 | 0.886 | -0.055 | 6884.1 | 0.000 |
| | | 12 | 0.876 | -0.113 | 7438.2 | 0.000 |
| | | 13 | 0.866 | 0.016 | 7980.0 | 0.000 |
| | | 14 | 0.856 | 0.069 | 8510.8 | 0.000 |
| | | 15 | 0.847 | -0.008 | 9031.3 | 0.000 |
| | | 16 | 0.838 | 0.019 | 9542.0 | 0.000 |
| | | 17 | 0.830 | 0.021 | 10043. | 0.000 |
| | | 18 | 0.822 | 0.006 | 10535. | 0.000 |
| | | 19 | 0.814 | 0.010 | 11018. | 0.000 |
| | | 20 | 0.807 | 0.006 | 11494. | 0.000 |
| | | 21 | 0.800 | -0.029 | 11963. | 0.000 |
| | | 22 | 0.792 | -0.045 | 12423. | 0.000 |
| | | 23 | 0.783 | -0.017 | 12873. | 0.000 |
| | | 24 | 0.774 | 0.002 | 13314. | 0.000 |
| | | 25 | 0.766 | 0.034 | 13745. | 0.000 |
| | | 26 | 0.758 | 0.031 | 14168. | 0.000 |
| | | 27 | 0.751 | 0.090 | 14585. | 0.000 |
| | | 28 | 0.746 | 0.012 | 14997. | 0.000 |
| | | 29 | 0.741 | 0.024 | 15404. | 0.000 |
| | | 30 | 0.738 | 0.000 | 15807. | 0.000 |
| | | 31 | 0.734 | -0.002 | 16207. | 0.000 |
| | | 32 | 0.732 | 0.061 | 16605. | 0.000 |
| | | 33 | 0.729 | 0.008 | 17001. | 0.000 |
| | | 34 | 0.727 | 0.017 | 17395. | 0.000 |

# DLog(Copper) Stationarity



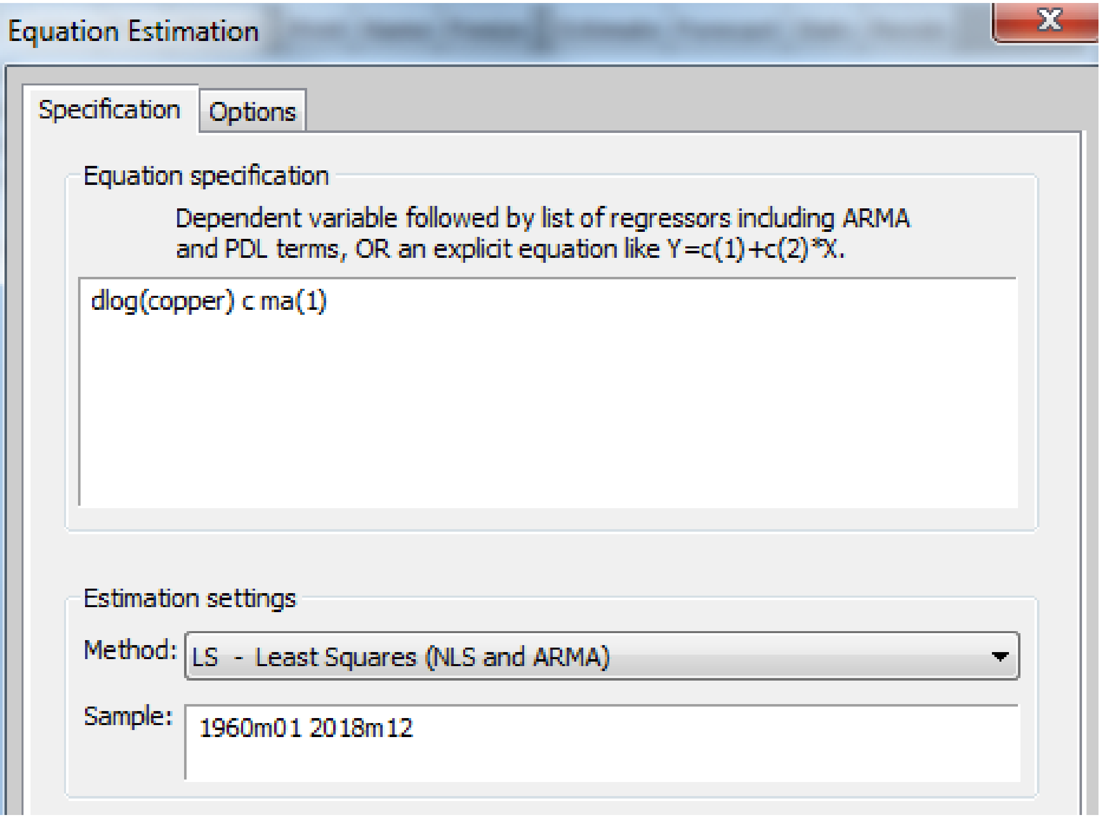
# DLOG(Copper) Example: Correlogram

Sample: 1960M01 2018M12
Included observations: 707

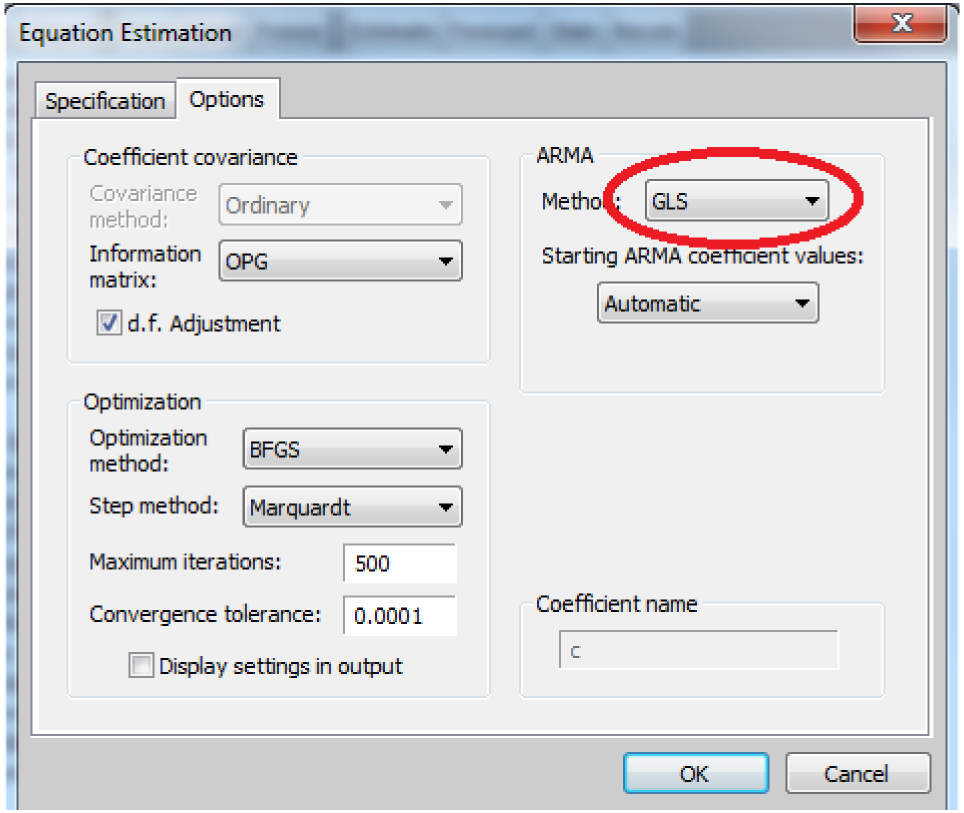
| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.331 | 0.331 | 77.832 | 0.000 |
| | | 2 | 0.011 | -0.110 | 77.924 | 0.000 |
| | | 3 | -0.018 | 0.016 | 78.155 | 0.000 |
| | | 4 | 0.010 | 0.015 | 78.231 | 0.000 |
| | | 5 | -0.004 | -0.017 | 78.241 | 0.000 |
| | | 6 | -0.006 | 0.003 | 78.265 | 0.000 |
| | | 7 | -0.054 | -0.060 | 80.362 | 0.000 |
| | | 8 | -0.117 | -0.090 | 90.141 | 0.000 |
| | | 9 | -0.077 | -0.011 | 94.397 | 0.000 |
| | | 10 | 0.021 | 0.046 | 94.708 | 0.000 |
| | | 11 | 0.084 | 0.063 | 99.849 | 0.000 |
| | | 12 | 0.072 | 0.029 | 103.55 | 0.000 |
| | | 13 | -0.010 | -0.043 | 103.63 | 0.000 |
| | | 14 | -0.067 | -0.053 | 106.87 | 0.000 |
| | | 15 | -0.073 | -0.049 | 110.68 | 0.000 |
| | | 16 | 0.014 | 0.041 | 110.82 | 0.000 |
| | | 17 | 0.005 | -0.028 | 110.84 | 0.000 |
| | | 18 | -0.107 | -0.103 | 119.16 | 0.000 |
| | | 19 | -0.081 | 0.012 | 123.94 | 0.000 |
| | | 20 | -0.016 | 0.012 | 124.14 | 0.000 |
| | | 21 | 0.023 | 0.014 | 124.52 | 0.000 |
| | | 22 | 0.013 | -0.022 | 124.65 | 0.000 |
| | | 23 | 0.057 | 0.046 | 127.02 | 0.000 |
| | | 24 | 0.023 | -0.011 | 127.41 | 0.000 |
| | | 25 | -0.012 | -0.007 | 127.52 | 0.000 |
| | | 26 | -0.009 | -0.011 | 127.58 | 0.000 |
| | | 27 | 0.007 | -0.006 | 127.62 | 0.000 |
| | | 28 | -0.046 | -0.061 | 129.18 | 0.000 |
| | | 29 | -0.079 | -0.037 | 133.82 | 0.000 |
| | | 30 | -0.061 | -0.010 | 136.58 | 0.000 |

- Even though stationary we can still model the residuals

# DLOG(Copper) Example: MA(1)



# OPTIONS: Generalised Least Squares - GLS



# DLOG(Copper) Example: MA(1)

Dependent Variable: DLOG(COPPER)
Method: ARMA Generalized Least Squares (Gauss-Newton)
Date: 01/23/19 Time: 15:48
Sample: 1960M02 2018M12
Included observations: 707
Convergence achieved after 6 iterations
Coefficient covariance computed using outer product of gradients
d.f. adjustment for standard errors & covariance

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.003041 | 0.003225 | 0.942853 | 0.3461 |
| MA(1) | 0.356941 | 0.035185 | 10.14475 | 0.0000 |

| | | | |
|---|---|---|---|
| R-squared | 0.119345 | Mean dependent var | 0.003026 |
| Adjusted R-squared | 0.118096 | S.D. dependent var | 0.067319 |
| S.E. of regression | 0.063219 | Akaike info criterion | -2.681416 |
| Sum squared resid | 2.817605 | Schwarz criterion | -2.668513 |
| Log likelihood | 949.8805 | Hannan-Quinn criter. | -2.676431 |
| F-statistic | 95.54071 | Durbin-Watson stat | 1.977221 |
| Prob(F-statistic) | 0.000000 | | |

# DLOG(Copper) Example: MA(1) – Residuals Corr.

Sample: 1960M01 2018M12
Included observations: 707
Q-statistic probabilities adjusted for 1 ARMA term

| Autocorrelation | Partial Correlation | | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|---|---|
| | | 1 | 0.011 | 0.011 | 0.0907 | |
| | | 2 | 0.018 | 0.018 | 0.3174 | 0.573 |
| | | 3 | -0.032 | -0.032 | 1.0338 | 0.596 |
| | | 4 | 0.025 | 0.025 | 1.4768 | 0.688 |
| | | 5 | -0.014 | -0.014 | 1.6202 | 0.805 |
| | | 6 | 0.008 | 0.006 | 1.6635 | 0.893 |
| | | 7 | -0.028 | -0.026 | 2.2283 | 0.898 |
| | | 8 | -0.091 | -0.093 | 8.2161 | 0.314 |
| | | 9 | -0.055 | -0.051 | 10.361 | 0.241 |
| | | 10 | 0.019 | 0.021 | 10.622 | 0.303 |
| | | 11 | 0.061 | 0.059 | 13.285 | 0.208 |
| | | 12 | 0.058 | 0.058 | 15.695 | 0.153 |
| | | 13 | -0.016 | -0.017 | 15.873 | 0.197 |
| | | 14 | -0.041 | -0.042 | 17.069 | 0.196 |
| | | 15 | -0.069 | -0.073 | 20.497 | 0.115 |
| | | 16 | 0.027 | 0.015 | 21.022 | 0.136 |
| | | 17 | 0.029 | 0.024 | 21.616 | 0.156 |
| | | 18 | -0.102 | -0.104 | 29.249 | 0.032 |
| | | 19 | -0.045 | -0.027 | 30.753 | 0.031 |
| | | 20 | -0.012 | 0.007 | 30.860 | 0.042 |
| | | 21 | 0.032 | 0.029 | 31.610 | 0.048 |
| | | 22 | -0.016 | -0.030 | 31.791 | 0.061 |
| | | 23 | 0.060 | 0.036 | 34.405 | 0.045 |
| | | 24 | 0.008 | 0.008 | 34.448 | 0.059 |

## ARIMA: Examples

| | |
|---|---|
| • ARIMA(2,0,1) | Eviews: *Y* c AR(1) AR(2) MA(1) |
| • ARIMA([4],0,0) | Eviews: *Y* c AR(4) |
| • ARIMA(0,0,2) | Eviews: *Y* c MA(1) MA(2) |
| • ARIMA(1,0,1) | Eviews: *Y* c AR(1) MA(1) |
| • ARIMA(1,1,1) | Eviews: D(*Y*) c AR(1) MA(1) |
| • ARIMA(1,2,1) | Eviews: D(*Y*, 2) c AR(1) MA(1) |

## Summary

- ARIMA modelling is an attempt to model the true data generating process

- Box-Jenkins Methodology - requires series to be stationary so use Dickey-Fuller tests and difference if necessary

  - Model Identification - ACF & PACF
  - Model Estimation - estimate several candidates

## EXTRA Readings


Box, G. E. P., and G. M. Jenkins, 1970, *Time Series Analysis, Forecasting and Control*. San Francisco: Holden Day.