Univariate Time Series Forecasting

Tomas Kana

4101 Macro Econometrics

ao.Univ.Prof. Dr. Michael Hauser

Mag.Dr. Thomas Url

Agenda

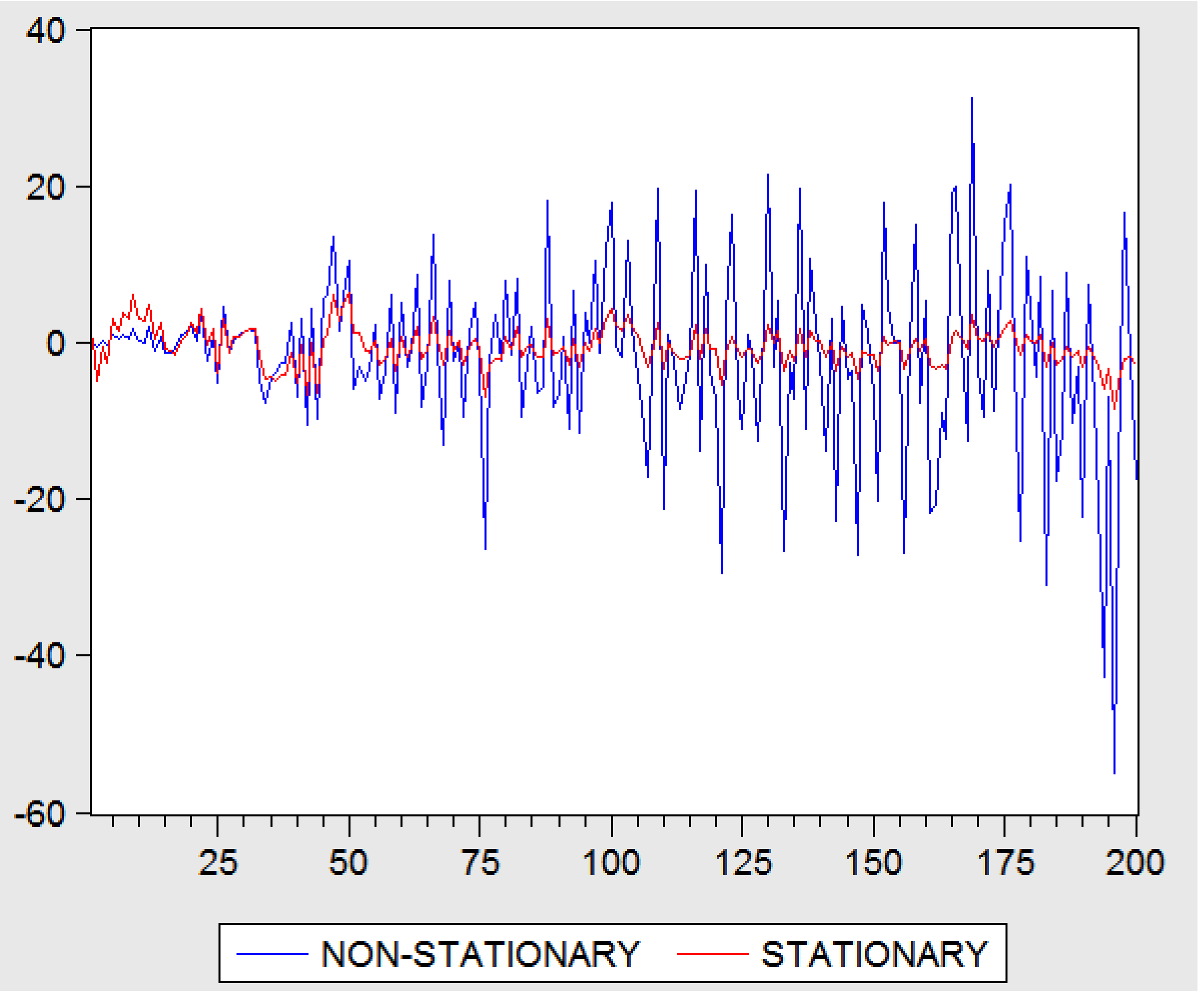
- Stationarity
- Autocorrelation/Autocovariance
- White Noise/Random Walk
- AR/MA/ARMA Models
- Model Estimation
- Forecasting
- Dickey-Fuller-Test

(Weak) Stationary Processes

- Stationary processes:
  1) Mean stationary: $\mu_t = \mu$
  2) Variance stationary: $\sigma_t^2 = \sigma^2$
  3) Covariance stationary: $\gamma_{t,s} = \gamma_s$

- Weak stationary: if 1), 2) and 3) are fulfilled

(Weak) Stationary Processes



Autocorrelation/Autocovariance

- Describes relationship between observations in different periods

Variance: $\gamma_0 = \sigma_y^2$

Autocovariance: $\gamma_s = E(y_t - \mu)(y_{t-s} - \mu)$

Autocorrelation: $\rho_s = \frac{\gamma_s}{\gamma_0}$

Autocovariance is symmetrical:

$$\gamma_s = Cov(y_{t-s}, y_t) = Cov(y_t, y_{t-s}) = Cov(y_{t+s}, y_t)$$

White Noise

- **Stationary process**

$$\boldsymbol{y_t = \epsilon_t}$$
$$\boldsymbol{E(\epsilon_t) = 0, V(\epsilon_t) = \sigma_\epsilon^2}$$

ACF: $\rho_0 = 1,\ \rho_1 = 0,\ \rho_2 = 0, \dots$

If $s \neq 0, \text{then}\ \gamma_s = 0$

Random Walk

- **Non-stationary process**
- Without drift: $\boldsymbol{y_t = y_{t-1} + \epsilon_t}$
- With drift: $\boldsymbol{y_t = c + y_{t-1} + \epsilon_t}$
- Example: $y_2 = c + y_1 + \epsilon_2$

$$y_2 = c + (c + y_0 + \epsilon_1) + \epsilon_2$$

$$y_2 = y_0 + 2c + \epsilon_1 + \epsilon_2$$

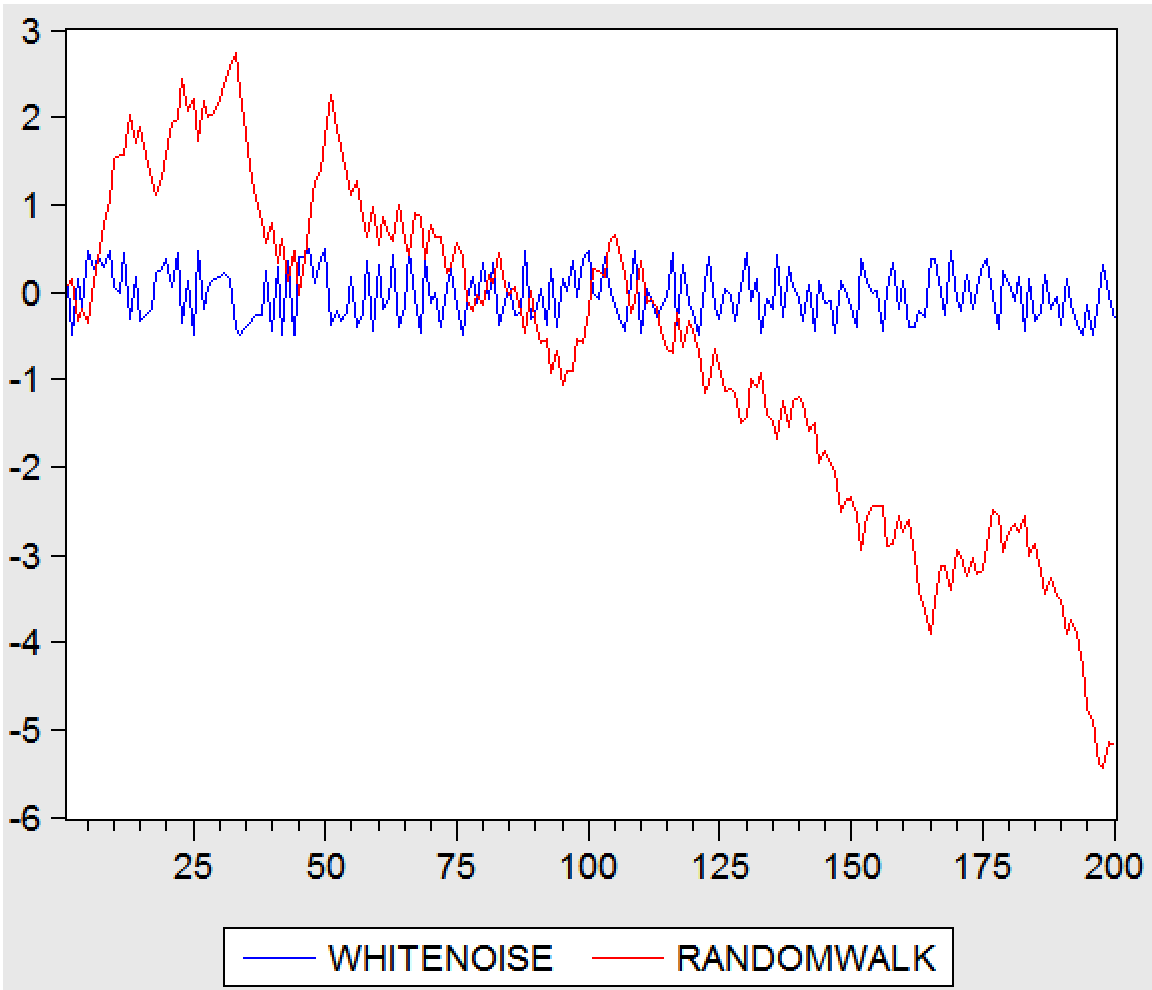
$$\boldsymbol{y_t = y_0 + ct + \sum_{j=1}^{t} \epsilon_j}$$

Random Walk

$$\boldsymbol{y_t = y_0 + ct + \sum_{j=1}^{t} \epsilon_j}$$

- Growing variance $(\sigma_t^2 = t\sigma_\epsilon^2)$
- Time-dependent expected value
- ACF: $\rho_0 = 1,\ \rho_1 = 1,\ \rho_2 = 1, \rho_3 = 1,\ldots$

White Noise / Random Walk



Transforming Non-Stationary Process into Stationary Process

- Differentiating: $y_t = x_t - x_{t-1}$
  - random walk
- Line fitting to observed values –> residuals from a linear trend as a new variable
  - deterministic trend
- Logarithm/square root – can stabilize the variance, if it is not constant
  - geometrical growth

AR Model

- Autoregression model: y_t depends on a weighted sum of the past values and a random shock
- AR(1): $y_t = \alpha y_{t-1} + \epsilon_t,\ s.t. |\alpha| < 1$
- Equiv. to: $y_t - \alpha y_{t-1} = \epsilon_t$

  s.t.: $\alpha < 1$

  ($\alpha = 1$ -> Random Walk!)

Lag Operator

- $Ly_t = y_{t-1}$
- $L^2 y_t = Ly_{t-1} = y_{t-2}$

- AR(1): $y_t - \alpha y_{t-1} = \epsilon_t$

  $(1 - \alpha L) y_t = \epsilon_t$

- AR(2): $y_t - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} = \epsilon_t$

  $\underbrace{(1 - \alpha_1 L - \alpha_2 L^2)}_{\text{„lag polynomial“}} y_t = \epsilon_t$

„lag polynomial“

MA Model

- Moving average model: observed value depends on a weighted linear sum of the past errors
- MA(1): $y_t = \mu + \epsilon_t + \beta \epsilon_{t-1}$
- or: $y_t = \mu + (1 - \beta L)\epsilon_t$

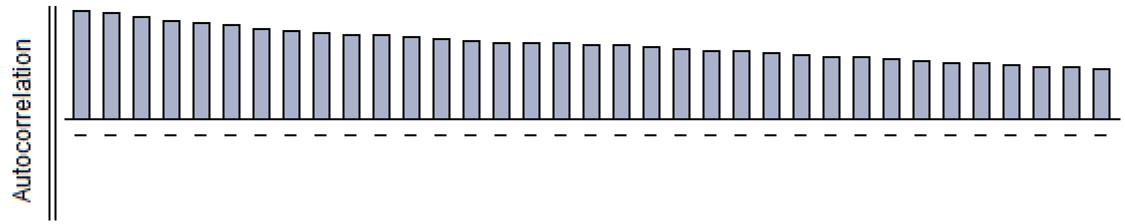
ARMA Model

- Combination of AR and MA processes

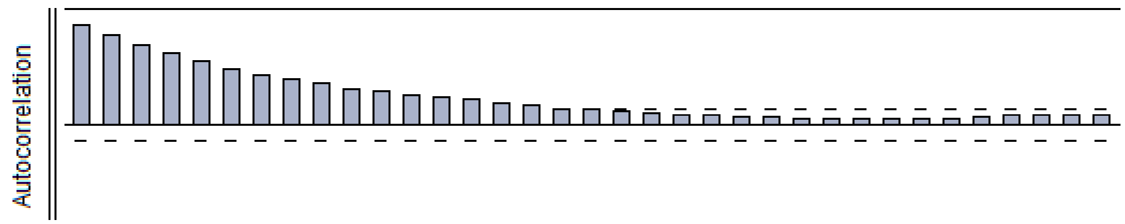
- ARMA(1,1): $y_t = \alpha y_{t-1} + \epsilon_t + \beta \epsilon_{t-1}$

ARMA Model

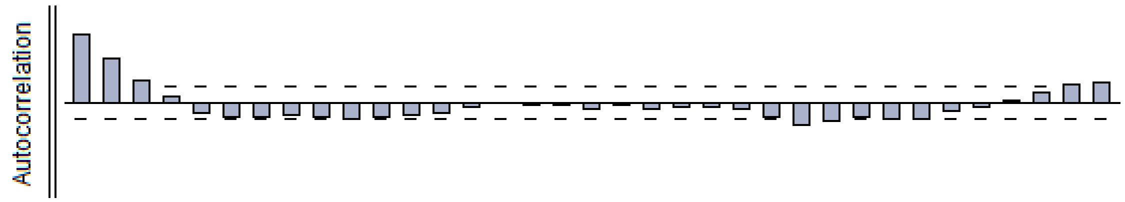
- ACF -> indicator of the correct model
  - Random Walk:



  - AR Process:



  - MA Process:



Estimation of the model

- Finding the parameters, which are most in accordance with the observed vlaues
- Unconditional least squares (ULS)
- Maximum likelihood (ML)

Forecasting with AR model

- AR(2):
  - $\hat{y}_{t+1} = \alpha_1 y_t + \alpha_2 y_{t-1} + \epsilon_{t+1}$
  - $\hat{y}_{t+2} = \alpha_1 \hat{y}_{t+1} + \alpha_2 y_t + \epsilon_{t+2}$
  - $\hat{y}_{t+3} = \alpha_1 \hat{y}_{t+2} + \alpha_2 \hat{y}_{t+1} + \epsilon_{t+3}$
  - etc.

- Convergence to mean

Forecasting with MA model

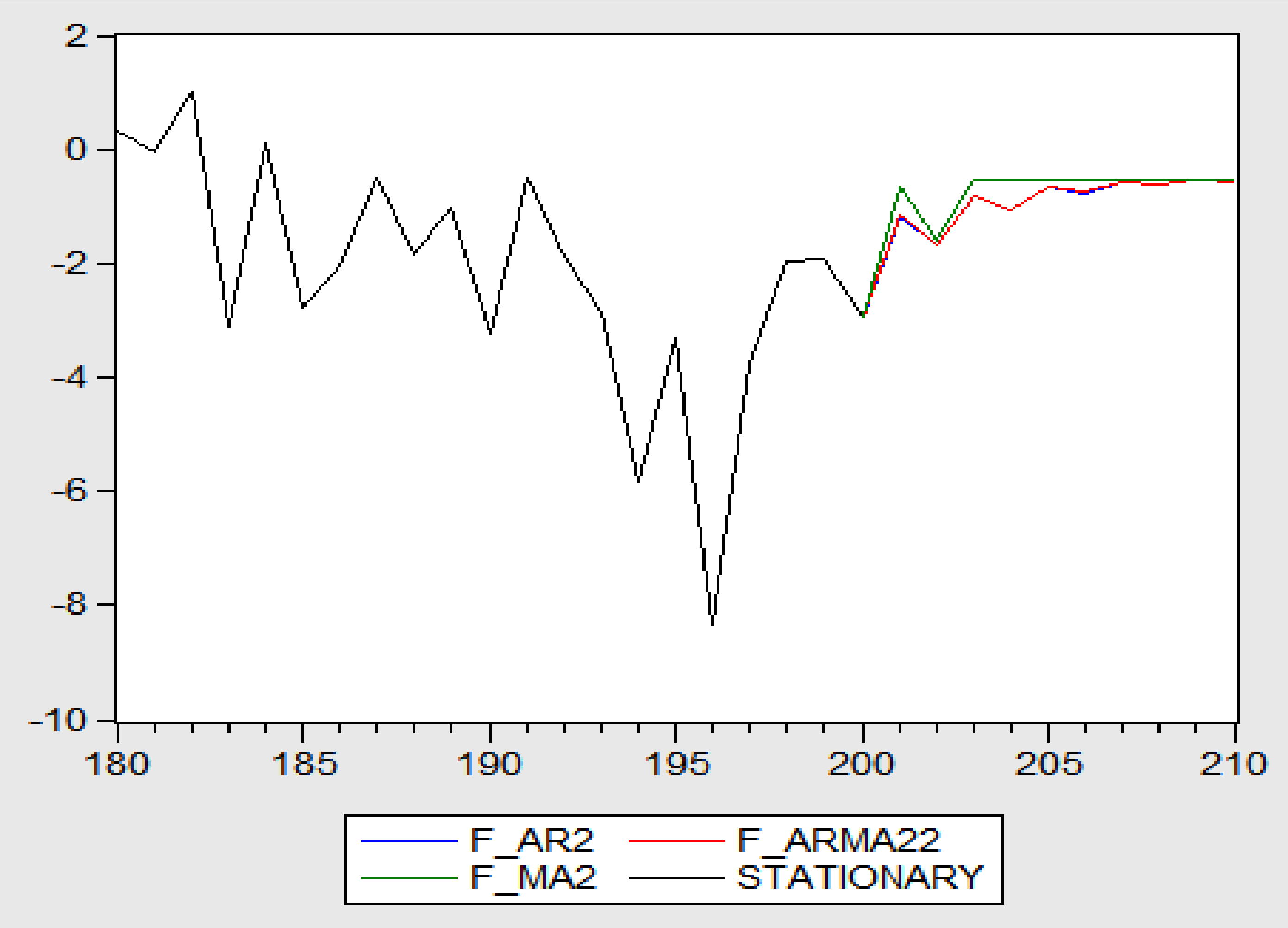
- MA(2):
  - $\hat{y}_{t+1} = \mu + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_t + \epsilon_{t+1}$
  - $\hat{y}_{t+2} = \mu + \beta_1 \epsilon_t + \beta_2 \epsilon_{t+1} + \epsilon_{t+2}$
  - $\hat{y}_{t+3} = \mu + \beta_1 \epsilon_{t+1} + \beta_2 \epsilon_{t+2} + \epsilon_{t+3}$
  - etc.

- MA(2) model -> forecast after 2 periods equal to mean

Forecasting with ARMA

- ARMA(2,2):
  - $\hat{y}_{t+1} = \mu + \alpha_1 y_t + \alpha_2 y_{t-1} + \beta_1 \epsilon_t + \beta_2 \epsilon_{t-1} + \epsilon_{t+1}$
  - $\hat{y}_{t+2} = \mu + \alpha_1 \hat{y}_{t+1} + \alpha_2 y_t + \beta_1 \epsilon_{t+1} + \beta_2 \epsilon_t + \epsilon_{t+2}$
  - $\hat{y}_{t+3} = \mu + \alpha_1 \hat{y}_{t+2} + \alpha_2 \hat{y}_{t+1} + \beta_1 \epsilon_{t+2} + \beta_2 \epsilon_{t+1} + \epsilon_{t+3}$
  - etc.

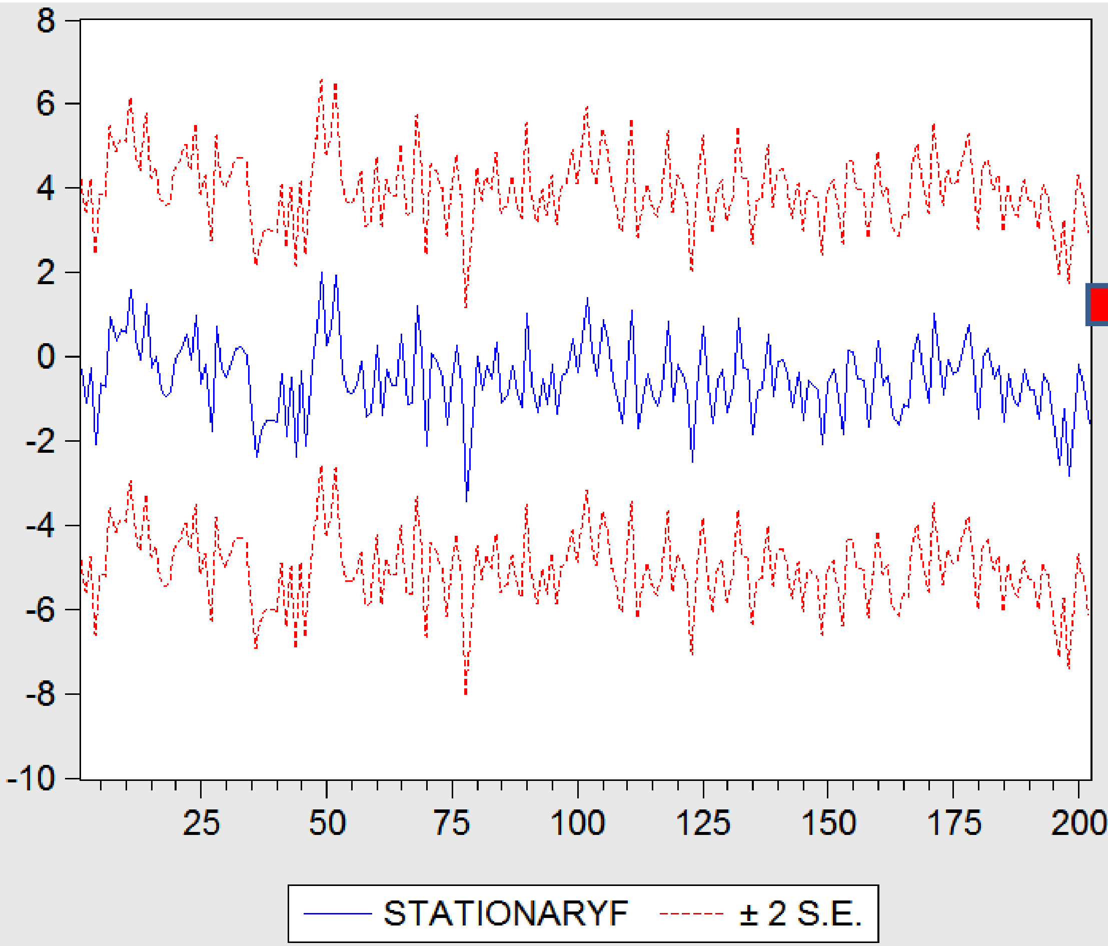
Forecast: AR(2) / MA(2) / ARMA(2,2)



Forecast Evaluation

- Root mean square error (RMSE)
- Measure of the differences between predicted and observed values (prediction errors)
- Measure of accuracy between different models
- $RMSE = \sqrt{\frac{\sum_{\tau=1}^{n}(y_{t+\tau}-\widehat{y_{t+\tau}})^2}{n}}$

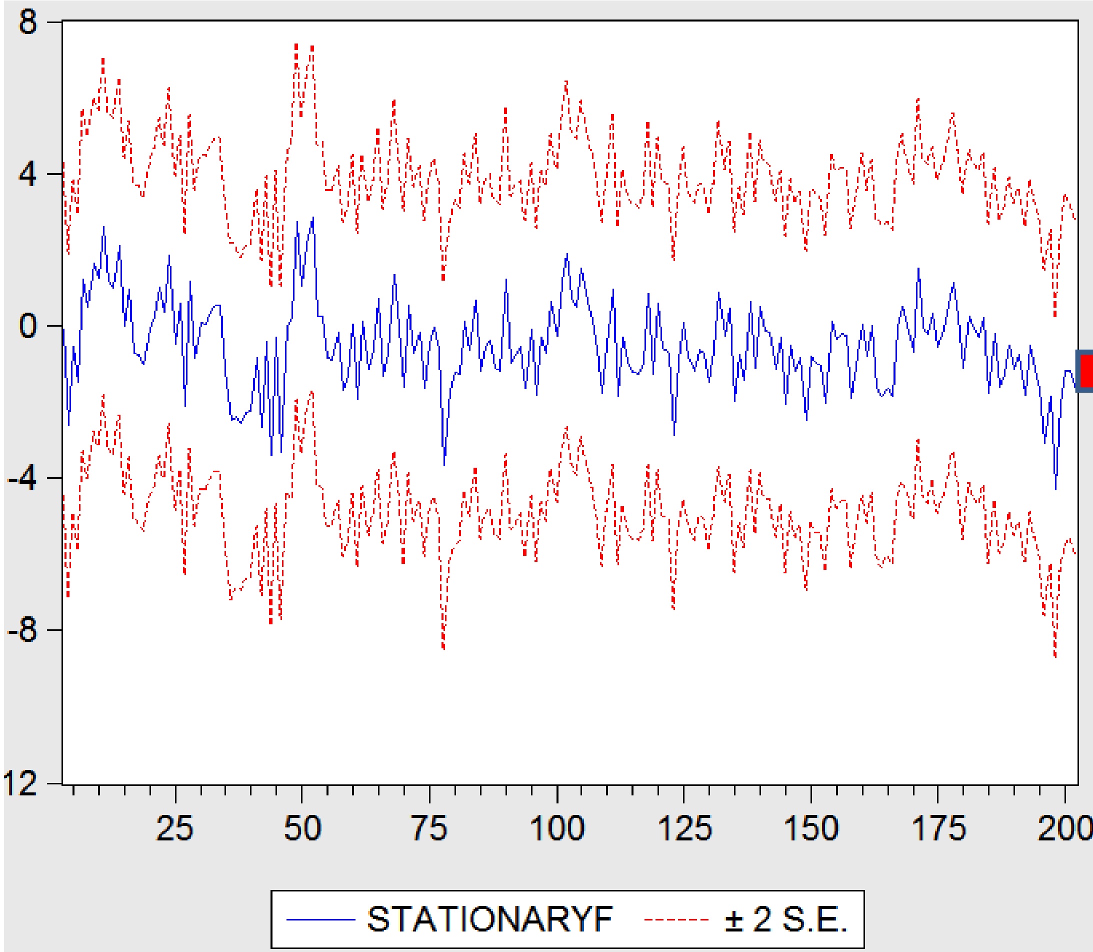
Forecast Evaluation – MA(2)



Forecast: STATIONARYF
Actual: STATIONARY
Forecast sample: 1 202
Included observations: 200

Root Mean Squared Error	2.229763
Mean Absolute Error	1.716068
Mean Abs. Percent Error	192.6628
Theil Inequality Coefficient	0.626694
Bias Proportion	0.000001
Variance Proportion	0.511601
Covariance Proportion	0.488398

Forecast Evaluation – ARMA(2,2)



Forecast: STATIONARYF
Actual: STATIONARY
Forecast sample: 1 300
Adjusted sample: 3 202
Included observations: 198

Root Mean Squared Error	2.160901
Mean Absolute Error	1.692509
Mean Abs. Percent Error	223.6721
Theil Inequality Coefficient	0.566507
Bias Proportion	0.000000
Variance Proportion	0.348334
Covariance Proportion	0.651666

Dickey-Fuller Test

- Unit root test
- Random Walk: $y_t = c + 1y_{t-1} + \epsilon_t$
- Stationary process: $y_t = c + \rho y_{t-1} + \epsilon_t$

$$\Delta y_t = (y_t - y_{t-1}) = c + \rho y_{t-1} + \epsilon_t - y_{t-1}$$
$$\boldsymbol{\Delta y_t = c + (\rho - 1)\, y_{t-1} + \epsilon_t}$$

- Dickey-Fuller Test:
- $H_0: \rho = 1, H_A: \rho < 1$
- $H_0: (\rho - 1) = 0, H_A: (\rho - 1) < 0$

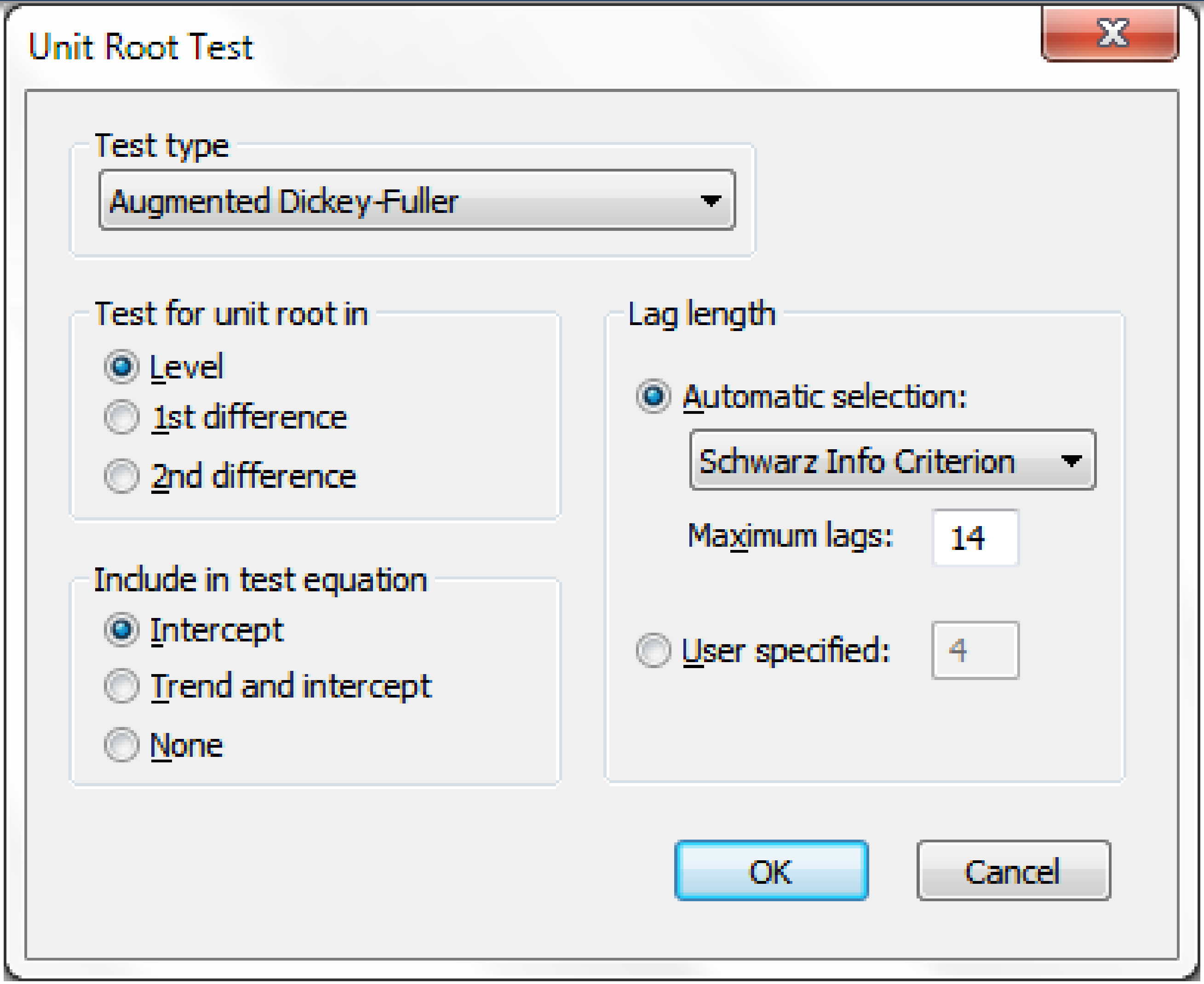
Dickey-Fuller Test

- $H_0: (\rho - 1) = 0,\ H_A: (\rho - 1) < 0$
- Regression of: $\Delta y_t = c + (\rho - 1) y_{t-1} + \epsilon_t$

$$DF = \frac{\widehat{\rho - 1}}{\sqrt{\widehat{Var}(\widehat{\rho - 1})}}$$

- Distribution dependent on drift (with/without) and number of regression coefficients

Dickey-Fuller Test



Null Hypothesis: STATIONARY has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic based on SIC, MAXLAG=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.150842	0.0000
Test critical values:	1% level	-3.463405	
	5% level	-2.875972	
	10% level	-2.574541	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(STATIONARY)
Method: Least Squares
Date: 04/20/13 Time: 10:23
Sample (adjusted): 3 200
Included observations: 198 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
STATIONARY(-1)	-0.380524	0.073876	-5.150842	0.0000
D(STATIONARY(-1))	-0.406209	0.064766	-6.271912	0.0000
C	-0.190071	0.155232	-1.224428	0.2223

R-squared	0.438667	Mean dependent var	0.010416
Adjusted R-squared	0.432910	S.D. dependent var	2.814306
S.E. of regression	2.119324	Akaike info criterion	4.355107
Sum squared resid	875.8493	Schwarz criterion	4.404929
Log likelihood	-428.1556	Hannan-Quinn criter.	4.375273
F-statistic	76.19376	Durbin-Watson stat	1.956004
Prob(F-statistic)	0.000000		