CHAPTER 3: MODELS WITH TRENDS (NONSTATIONARY MODELS)

Deterministic Trends (Drift Model)

- Suppose that a series changes by the same amount for every period

$$y_t = y_{t-1} + a_0 \qquad (\text{or } \Delta y_t = a_0)$$

where a_0 is the drift term. The general solution for y_t is given by

$$y_t = y_0 + a_0 t$$

- That is, y_t follows a time trend beginning from y_0
- This is a nonstationary process. This is because neither its mean, nor its variance is constant and time invariant

Stochastic Trend (Random Walk Model)

- Sequence $\{y_t\}$ follows a random walk process if

$$y_t = y_{t-1} + \varepsilon_t \qquad (\text{or } \Delta y_t = \varepsilon_t)$$

- Notice that unlike deterministic trend, changes in y_t are unpredictable at time t. In this case, we say y_t has stochastic trend, as all stochastic shocks have nondecaying effects on the $\{y_t\}$ sequence
- The general solution to the random walk model is

$$y_t = y_0 + \sum_{i=1}^{t} \varepsilon_i$$

- The last observation is the unbiased estimator of all future values of y_{t+s}

$$E_t y_{t+s} = y_t + E_t \sum_{i=1}^{s} \varepsilon_{t+i} = y_t$$

- The expectation of the process is not time invariant

$$E_o y_t = y_o$$

- Moreover, the variance is not constant but depends on t

$$var(y_t) = var(\varepsilon_t + \varepsilon_{t-1} + ... + \varepsilon_1) = t\sigma^2$$

- Thus, the random walk process is nonstationary
- As $t \to \infty$, the variance of y_t also approaches infinity (this is intuitive as we wouldn't know where $\{y_t\}$ sequence might go)
- Since the memory of a random walk process is very high (the coefficient of the autoregressive part, y_{t-1}, is 1 in the model), the autocorrelation function for a random walk process (ρ_s) is close to unity for small s

The Random Walk Plus Drift Model

- This model augments the random walk model by adding a constant term a_0, so that

$$y_t = y_{t-1} + a_0 + \varepsilon_t \qquad (\text{or } \Delta y_t = a_0 + \varepsilon_t)$$

- The general solution for y_t is given by

$$y_t = y_0 + a_0 t + \sum_{i=1}^{t} \varepsilon_i \tag{4.5}$$

- There are two nonstationary components
  - a linear deterministic trend
  - the stochastic trend $\sum \varepsilon_i$

REMOVING THE TREND

- A reason for trying to stationarize a time series is to be able to obtain meaningful sample statistics such as means, variances, and correlations with other variables
- Moreover, using non-stationary time series data produces unreliable and spurious results and leads to poor understanding and forecasting (law of large numbers, central limit theorem hold for stationary random variables)
- Below, we analyze methods for making a series stationary, appropriateness of which depend on whether the trend has a deterministic, or a stochastic component

Differencing

- First consider the solution for the random walk plus drift model

$$y_t = y_{t-1} + a_0 + \varepsilon_t$$

  Taking the first difference, we obtain

$$\Delta y_t = a_0 + \varepsilon_t \tag{1}$$

  Clearly, the Δy_t sequence—equal to a constant plus a white-noise disturbance—is stationary

- The above process is called ARIMA (0,1,0) model (ARIMA: autoregressive integrated moving average)
- The model has no AR and MA component, but requires first differencing to be stationary
- Now consider the general model ARIMA model

$$A(L)y_t = B(L)\varepsilon_t \tag{2}$$

  where $A(L)$ and $B(L)$ are polynomials of orders p and q in the lag operator L

- We know that stationary of y_t requires all all roots of $A(L)$ lie outside the unit circle
- Suppose this condition is not satisfied and y_t has a single unit root

- Then we can factor $A(L)$ into two components $(1-L)A^*(L)$

$$A(L) = (1-L)A^*(L),$$

where $A^*(L)$ is a polynomial of order $p-1$. Since $A(L)$ has only one unit root, it follows that all roots of $A^*(L)$ are outside the unit circle

- Thus, we can write (2) as

$$(1-L)A^*(L)y_t = B(L)\varepsilon_t$$

so that

$$A^*(L)\Delta y_t = B(L)\varepsilon_t$$

- Notice that as all roots of $A^*(L)$ lies outside the unit circle, the Δy_t sequence is stationary
- Hence, differencing the data transforms an I(1) process to an I(0)
- *The general point is that the d^{th} difference of a process with d unit roots is stationary*
- These models are called *Difference Stationary Models*
- Such a sequence is integrated of order d and denoted by I(d)

Detrending

- Not all nonstationary models can be transformed into well-behaved ARMA models by appropriate differencing
- Consider, for example, a model that is the sum of a deterministic trend and a pure noise component:

$$y_t = y_0 + a_1 t + \varepsilon_t$$

- The first difference of y_t is not well-behaved because

$$\Delta y_t = a_1 + \varepsilon_t - \varepsilon_{t-1}$$

- Here, Δy_t is not invertible (it cannot be expressed in the form of an convergent or finite autoregressive process)
- Instead, an appropriate way to transform this model is to estimate the regression

$$y_t = a_0 + a_1 t + e_t,$$

where e_t is the estimated values of the ε_t series, and a_0 is the estimated value of the y_0

- Simply subtracting the estimated values of the y_t sequence from the actual values (detrending) yields an estimate of the stationary sequence e_t
- Hence, the above model is called *Trend Stationary Model*
- The detrended (or differenced) processes can then be modeled using traditional methods (such as ARMA estimation)
- In general, detrending is accomplished by regressing y_t on a deterministic polynomial time trend

$$y_t = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n + e_t$$

- To find the appropriate degree of the polynomial the common practice is to estimate the regression equation using the large of n, if the t-statistics indicates a_n is zero, consider a polynomial trend of order $n-1$, and continue this way until a nonzero coefficient is found
- Though computer programs give you the appropriate lag length chosen according to AIC or SBC criterias

The Effect of a Unit Root on Regression Residuals

- Consider the regression equation

$$y_t = a_0 + a_1 z_t + e_t \tag{4.12}$$

- The assumptions of the classical regression model necessitate that both the y_t and z_t sequences be stationary and that the errors have a zero mean and a finite variance
- In the presence of nonstationary variables, there might be what Granger and Newbold (1974) call a spurious regression
- A spurious regression has a high R^2 and t-statistics that appear to be significant, but the results are without any economic meaning.
- The regression output "looks good," but the least-squares estimates are not consistent and the customary tests of statistical inference do not hold
- Let's generate two sequences,y_t and z_t, as independent random walks using the formulas:

$$y_t = 0.2 + y_{t-1} + \varepsilon_{yt}$$

$$z_t = -0.1 + z_{t-1} + \varepsilon_{zt}$$

  where ε_{yt} and ε_{zt} are white-noise processes that are independent of each other
- Since the y_t and z_t sequences are independent of each other, (4.12) is necessarily meaningless; any relationship between the two variables is spurious
- Surprisingly, Granger and Newbold (1974) were able to reject the null hypothesis $a_1 = 0$ in approximately 75% of the cases
  - The reason is that the deterministic drift terms that cause the sustained increase in y_t and the overall decline in z_t. Hence, it appears that the two series are inversely related to each other and the correlation between the variables gets close to 1
  - Moreover, the estimated residuals from a spurious regression (e_t) will exhibit a high degree of autocorrelation, and its the variance becomes infinitely large as t increases, which is inconsistent with the distributional theory underlying the use of OLS
- *The Results:*
- If the y_t and z_t sequences are integrated of different orders, regression equations using such variables are meaningless
- If the nonstationary y_t and z_t sequences are integrated of the same order, and if the residual sequence contains a stochastic trend, the regression is spurious

- In this case, it is often recommended that the regression equation be estimated in first differences

$$\Delta y_t = a_1 \Delta z_t + \Delta e_t$$

- If the nonstationary y_t and z_t sequences are integrated of the same order, and if the residual sequence is stationary, y_t and z_t are *cointegrated*
  - A trivial example of a cointegrated system occurs if ε_{yt} and ε_{zt} are perfectly correlated

Dickey-Fuller Tests

- This section outlines a procedure to determine whether $a_1 = 1$ in the model

$$y_t = a_1 y_{t-1} + e_t$$

- Begin by subtracting y_{t-1} from each side of the equation in order to write the equivalent form

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

where $\gamma = a_1 - 1$

- Testing the hypothesis $a_1 = 1$ is equivalent to testing the hypothesis $\gamma = 0$
- *Example:* Suppose the estimate of

$$y_t = a_1 y_{t-1} + e_t$$

such that $a_1 = 0.9546$ with a standard error of 0.030

- The OLS regression in the form

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

will yield an estimate of γ equal to -0.0454 with the same standard error of 0.030

- Hence, the associated t-statistics for the hypothesis $\gamma = 0$ is $-0.0454/0.03 = -1.5133$
- Dickey and Fuller (1979) consider three different regression equations that can be used to test for the presence of a unit root:

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \varepsilon_t$$

- The parameter of interest in all the regression equations is γ; if $\gamma = 0$, the y_t sequence contains a unit root
- In their Monte Carlo study, Dickey and Fuller (1979) found that the critical values for $\gamma = 0$ depend on the form of the regression (whether an intercept and/or time trend is included in the regression equation) and sample size

- These critical values are unchanged when the last equations above are replaced by the autoregressive processes:

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t$$

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t$$

- Tests including lagged changes are called *Augmented Dickey–Fuller tests*
- Consider the pth order autoregressive process

$$y_t = a_0 + a_1 y_{t-1} + ... + a_p y_{t-p} + \varepsilon_t$$

  we do not get into details but this equation can be written as

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \varepsilon_t \tag{4.30}$$

  where $\gamma = -(1 - \sum_{i=1}^{p} a_i)$ and $\beta_i = \sum_{i=1}^{p} a_j$

- Here, if $\sum a_i = 1$ so that $\gamma = 0$, and the system has a unit root
- Clearly, the simple regression

$$\Delta y_t = a_0 + \gamma y_{t-1} + \varepsilon_t$$

  is inadequate to this task if (4.30) is the true data-generating process

  – We cannot properly estimate γ and its standard error unless all of the autoregressive terms are included in the estimating equation

- Since the true order of the autoregressive process is unknown, we need to select the appropriate lag length

  – Including too many lags reduces the power of the test and a loss of degrees of freedom

  – One approach is to use the general-to-specific methodology. Start with a relatively long lag length and pare down the model by the usual t-test and/or F-tests, or by AIC an SBC tests

Example: The file HW2_data.xlsx contains the U.S. data from column F onwards. Use the real GDP data in the file

a-) Form the log of real GDP as $ly_t = log(RGDP)$. Form the autocorrelations. By using augmented Dickey–Fuller test check if the series is stationary

Answer:

```
cd "E:\Dropbox\teaching\2015-16 TOBB-ETU\IKT 553 Applied Time Series 2"
wfcreate question q 1960Q1 2012Q4
import "HW2_data.xlsx" range="Sayfa1!F2:J213"
rename series02 Tbill
rename series03 r5
rename series04 RGDP
rename series05 Potent
series ly=@log(RGDP)
ly.correl(10)
```

Date: 01/08/16 Time: 09:51
Sample: 1960Q1 2012Q4
Included observations: 212

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA

```
ly.uroot(adf,info=aic)
```

b-) Form the log of real GDP as $ly_t = log(RGDP)$. Detrend the data with a linear time trend and obtain the residuals. Then check if the residual series is stationary

Answer:

```
ls ly c @trend
resid.correl(10)
```

Date: 01/08/16 Time: 09:52
Sample: 1960Q1 2012Q4
Included observations: 212

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA

```
resid.uroot(adf,info=aic)
```

c-) Find the growth rate of real GDP as $ly_t - ly_{t-1}$. Then check if the series is stationary.

Answer:

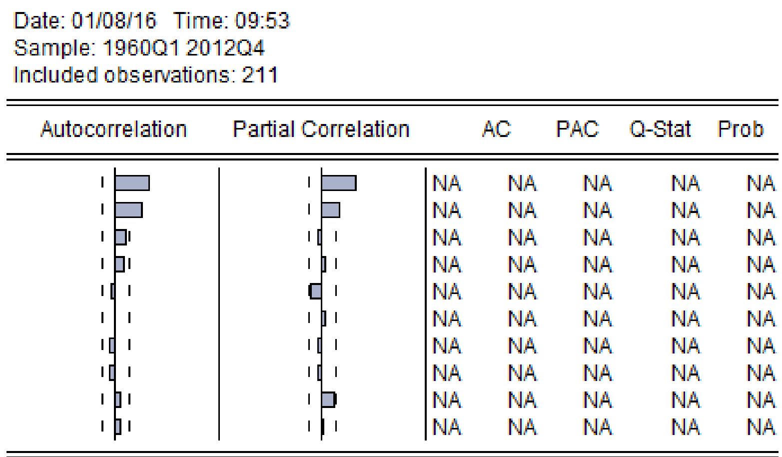
```
series y_gr=ly-ly(-1)
```

y_gr.uroot(adf,info=aic)

d-) Using ACF and PACF, and also the Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC) model selection criterias, find the parsimonous model that best represent the growth rate of real GDP data (find the ARMA representation of the data). Compare your result with the lag length selected automatically in the unit root test in part c-).

Answer: **-**AR(1)-

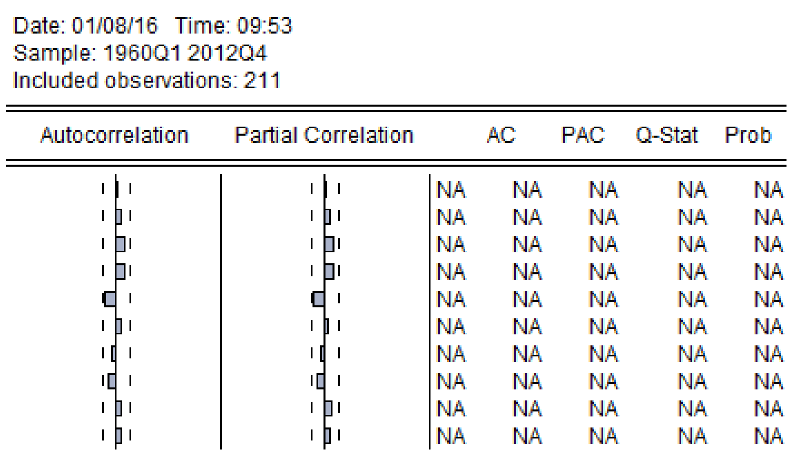
y_gr.correl(10)


Date: 01/08/16 Time: 09:53
Sample: 1960Q1 2012Q4
Included observations: 211

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA

ls y_gr c ma(1) ma(2)

resid.correl(10)


Date: 01/08/16 Time: 09:53
Sample: 1960Q1 2012Q4
Included observations: 211

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA
		NA	NA	NA	NA	NA

ls y_gr c ar(1)

Dependent Variable: Y_GR
Method: Least Squares
Date: 01/28/16 Time: 13:12
Sample (adjusted): 1960Q3 2012Q4
Included observations: 210 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007509	0.000857	8.764574	0.0000
AR(1)	0.349344	0.064718	5.397922	0.0000

R-squared	0.122872	Mean dependent var	0.007495
Adjusted R-squared	0.118655	S.D. dependent var	0.008605
S.E. of regression	0.008078	Akaike info criterion	-6.789831
Sum squared resid	0.013573	Schwarz criterion	-6.757953
Log likelihood	714.9322	Hannan-Quinn criter.	-6.776944
F-statistic	29.13756	Durbin-Watson stat	2.123387
Prob(F-statistic)	0.000000		

Inverted AR Roots	.35

resid.correl(10)

Date: 01/28/16 Time: 13:10
Sample: 1960Q1 2012Q4
Included observations: 210

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.063	-0.063	0.8563	0.355
		2	0.157	0.153	6.1057	0.047
		3	-0.007	0.011	6.1177	0.106
		4	0.114	0.093	8.9195	0.063
		5	-0.098	-0.091	11.021	0.051
		6	0.052	0.013	11.613	0.071
		7	-0.047	-0.020	12.103	0.097
		8	-0.066	-0.090	13.059	0.110
		9	0.064	0.088	13.955	0.124
		10	0.025	0.042	14.090	0.169