Econometrics

Slides-07: Nonstationary Processes Identification, Testing and Estimation

Dr. Rachida Ouysse
School of Economics[1]

Plan.

1. Stochastic and Deterministic Non-stationary Processes
   - Properties of Deterministic Non-stationary Process
   - Properties of Stochastic Non-stationary Process
     1. Random Walk Process
     2. Random Walk Process with a drift
   - Transformation to achieve Stationarity
2. Unit Root Tests
   1. Dickey Fuller Test: Basic
   2. Dickey Fuller Test: Intercept
   3. Dickey Fuller Test: Intercept and Trend
   4. Augmented Dickey-Fuller Test
3. Power consideration
4. Selection of model for testing

Stationarity versus Non-stationarity

A series y_t is cova[illegible]weakly **stationary** if it satisfies:

- $E\left(y_t\right) = \mu <$ [illegible]
- $Var\left(y_t\right) = E\left(y_t - \mu\right)^2 = \sigma^2 < \infty$
- $Cov\left(y_t, y_{t-k}\right) = E\left(y_t - \mu\right)\left(y_{t-k} - \mu\right) = \gamma_k \quad \forall k$

$\rightarrow$ The first and the second moment of the distribution of y_t are not affected by an arbitrary shift along the time axis.

We can distinguish two types of **non-stationarity**:

- Deterministic non-stationarity
- Stochastic non-stationarity

Deterministic Non-stationarity Process

Deterministic non-s

Consider a model wi end:

$$y_t = \alpha_0 + \phi t + \varepsilon_t$$

where t is a linear trend, i.e. $t = 1, 2, 3, \ldots$, and ε_t is a white noise disturbance term.

$$\rightarrow \quad E(y_t) = \alpha_0 + \phi t = \alpha_0 + \sum_{i=1}^{t} \phi$$

$\rightarrow$ mean is time-varying

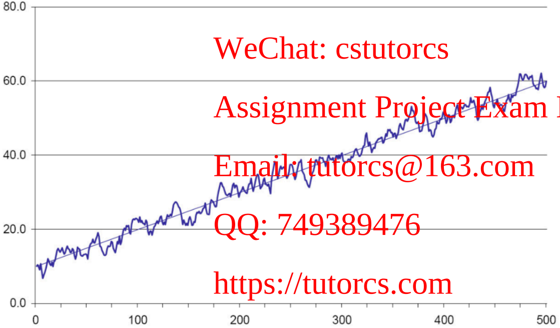
This model can be extended with AR and MA terms, e.g.:

$$y_t = \alpha_0 + \phi t + \alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t$$

$$\rightarrow \quad E(y_t) = (\alpha_0 + \phi t) / (1 - \alpha_1) \quad \text{if} \quad |\alpha_1| < 1$$

Deterministic Non-Stationary Process

Figure 2 : Model with linea... 00):
$y_t = 2 + 0.02t + 0.8y_{t-1} + ($...



Deterministic Non-Stationary Process

Consider a model with a ... eak:

$$y \ldots + \varepsilon_t$$

where $D_t = 0$ for $t = 1, \ldots, T_1$; $D_t = 1$ for $t = T_1 + 1, \ldots, T$, and ε_t is a white noise disturbance term.

$$\rightarrow \quad E(y_t) = \alpha_0 \quad \text{for} \quad t=1,\ldots,T_1$$
$$E(y_t) = \alpha_0 + \phi \quad \text{for} \quad t=T_1+1,\ldots,T$$

$\rightarrow$ mean is time-varying

This model can be extended with AR and MA terms, e.g.:

$$y_t = \alpha_0 + \phi D_t + \alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t$$
$$\rightarrow \quad E(y_t) = \alpha_0 / (1 - \alpha_1) \quad \text{for} \quad t=1,\ldots,T_1 \quad \text{if} \quad |\alpha_1|<1$$
$$E(y_t) = (\alpha_0 + \phi) / (1 - \alpha_1) \quad \text{for} \quad t=T_1+1,\ldots,T \quad \text{if} \quad |\alpha_1|<1$$

Deterministic Non-Stationary Process

Figure 3 : Model with stru = 100):
$y_t = 2 + 2D_t + 0.8y_{t-1} + 0.$

Stochastic Non-Stationary Process

Stochastic non-stationarity

Consider the AR(1) process:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t$$

The MA representation is given by:

$$y_t = \alpha_0 \sum_{i=0}^{\infty} \alpha_1^i + \alpha_1^{\infty} y_0 + \sum_{i=0}^{\infty} \alpha_1^i \varepsilon_{t-i}$$

Depending on the value for α_1, two cases can be distinguished:

- Stationary case: $|\alpha_1| < 1 \quad \Rightarrow \alpha_1^i \to 0 \quad \text{as} \quad i \to \infty$
  $\to$ shocks gradually die out
- Unit root case: $|\alpha_1| = 1 \quad \Rightarrow \alpha_1^i = 1 \quad \forall i$
  $\to$ shocks persist in the system.
- Explosive case: $|\alpha_1| > 1 \quad \Rightarrow \alpha_1^i \to \infty \quad \text{as} \quad i \to \infty$
  $\to$ shocks have an increasingly large influence.

Stochastic Non-Stationary Process

Both in the unit root and [illegible] case the mean, variance, covariances, . . . , are time[illegible]
$\rightarrow$ the series is said to exhibit **stochastic non-stationarity**.

As the explosive case does not describe many data series, typically $|\alpha_1| > 1$ is ignored and $\alpha_1 = 1$ is used to describe stochastic non-stationarity.

Now reconsider the unit root case:

$$y_t = \alpha_0 + y_{t-1} + \varepsilon_t$$

Again, two cases can be distinguished:

- Random walk: $\alpha_0 = 0$
- Random walk with drift: $\alpha_0 \neq 0$

Random Walk Process

The random walk model

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

Backward iterating the process yields:

$$\begin{aligned} y_t &= y_{t-1} + \varepsilon_t \\ &= y_{t-2} + \varepsilon_t + \varepsilon_{t-1} \\ &= y_{t-3} + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} \\ &\cdots \\ &= y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i} \end{aligned}$$

The unit root in the AR process causes the series to exhibit a **stochastic trend**, i.e. $\sum_{i=0}^{t-1} \varepsilon_{t-i}$.

Random Walk Process

Properties of a random walk

- Expected value

$$E_t(y_t) = E\left(y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right) = y_t$$

$$E_t(y_{t+1}) = E_t(y_t + \varepsilon_{t+1}) = y_t$$

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

- Variance

$$var(y_t) = var\left(y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right) = E\left(\sum_{i=0}^{t-1} \varepsilon_{t-i}\right)^2 = t\sigma^2$$

- Covariance

$$cov(y_t, y_{t-s}) = E\left((\varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1)(\varepsilon_{t-s} + \varepsilon_{t-s-1} + \dots + \varepsilon_1)\right)$$
$$= E\left(\varepsilon_{t-s}^2 + \varepsilon_{t-s-1}^2 + \dots + \varepsilon_1^2\right) = (t-s)\sigma^2$$

$\rightarrow$ The random walk meanders without exhibiting any tendency to increase or decrease.

Random Walk Process

Properties of a random walk

- Expected value

$$E_t(y_t) = E\left(y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right) = y_t$$

$$E_t(y_{t+1}) = E_t(y_t + \varepsilon_{t+1}) = y_t$$

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

- Variance

$$var(y_t) = var\left(y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right) = E\left(\sum_{i=0}^{t-1} \varepsilon_{t-i}\right)^2 = t\sigma^2$$

- Covariance

$$\begin{aligned} cov(y_t, y_{t-s}) &= E\left((\varepsilon_t + \varepsilon_{t-1} + \ldots + \varepsilon_1)(\varepsilon_{t-s} + \varepsilon_{t-s-1} + \ldots + \varepsilon_1)\right) \\ &= E\left(\varepsilon_{t-s}^2 + \varepsilon_{t-s-1}^2 + \ldots + \varepsilon_1^2\right) = (t-s)\sigma^2 \end{aligned}$$

$\rightarrow$ The random walk meanders without exhibiting any tendency to increase or decrease.

Random Walk: Simulated example

Figure 4 : Random walk $y_t = y_{t-1} + \varepsilon_t$

Random Walk: Simulated example

Figure 6 : Random wal $y_t = y_{t-1} + \varepsilon_t$

Random Walk with a Drift

The random walk with drift is given by:

$$y_t = \alpha_0 + y_{t-1} + \varepsilon_t$$

Backward iterating the process yields:

$$\begin{aligned} y_t &= \alpha_0 + y_{t-1} + \varepsilon_t \\ &= \alpha_0 + \alpha_0 + y_{t-2} + \varepsilon_t + \varepsilon_{t-1} \\ &= \alpha_0 + \alpha_0 + \alpha_0 + y_{t-3} + \varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2} \\ &\cdots \\ &= y_0 + \alpha_0 t + \sum_{i=0}^{t-1} \varepsilon_{t-i} \end{aligned}$$

The intercept in the non-stationary AR model adds a **deterministic trend**, i.e. $\alpha_0 t$, to the stochastic trend $\sum_{i=0}^{t-1} \varepsilon_{t-i}$.

Random Walk with a Drift: Properties

Properties of a random w

- Expected value

$$E_t(y_t) = E\left(y_0 + \alpha_0 t + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right) = y_t$$

$$E_t(y_{t+s}) = E_t\left(\sum_{i=1}^{s} \alpha_0 + y_t + \sum_{i=1}^{s} \varepsilon_{t+i}\right) = y_t + s\alpha_0$$

- Variance

$$var(y_t) = var\left(y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i}\right) = E\left(\sum_{i=0}^{t-1} \varepsilon_{t-i}\right)^2 = t\sigma^2$$

- Covariance

$$cov(y_t, y_{t-s}) = E((\varepsilon_t + \varepsilon_{t-1} + \dots + \varepsilon_1)(\varepsilon_{t-s} + \varepsilon_{t-s-1} + \dots + \varepsilon_1))$$
$$= E(\varepsilon_{t-s}^2 + \varepsilon_{t-s-1}^2 + \dots + \varepsilon_1^2) = (t-s)\sigma^2$$

The random walk with drift meanders with a tendency to increase ($\alpha_0 > 0$) or to decrease ($\alpha_0 < 0$).

Random Walk with a Drift: Properties

Figure 8 : Random walk w 500): $y_t = 0.1 + y_{t-1} + \varepsilon_t$

Transformation to Stationarity

Deterministic and stochastic non-stationary series require different treatments to induce stationarity.

Deterministic non-stationarity
Consider the model:

$$y_t = \alpha_0 + \phi t + \alpha_1 y_{t-1} + \epsilon_t$$

with $|\alpha_1| < 1$.

- Detrend y_t by subtracting $\frac{\phi}{1-\alpha_1} t$
- $y_t - \frac{\phi}{1-\alpha_1} t$ is stationary
- y_t is said to be **trend-stationary**, i.e. y_t is stationary around a linear trend

Transformation to Stationarity

Stochastic non-stationar

Consider the model:

$$y_t = \alpha_0 + y_{t-1} + \varepsilon_t$$

- Detrend y_t by taking a first-difference transformation:

$$y_t - y_{t-1} = \alpha_0 + \varepsilon_t$$
$$\Delta y_t = \alpha_0 + \varepsilon_t$$

- Δy_t is stationary

- y_t is said to be **difference-stationary**, i.e. y_t is stationary after taking first-differences

Unit Root Tests

How do we test for a

- **Examine the AC**

  Problem: it can be shown that in the random walk model, the autocorrelation coefficients are given by:

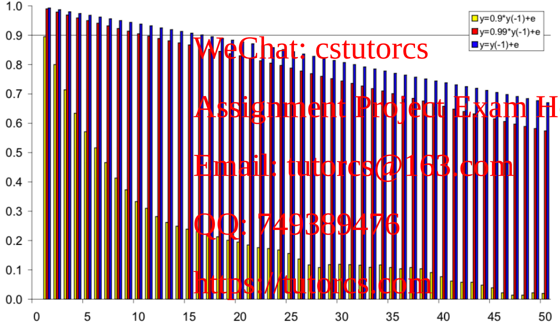
$$\rho_k = \frac{cov\left(y_t, y_{t-k}\right)}{\sqrt{var\left(y_t\right) var\left(y_{t-k}\right)}} = \frac{(t-k)\,\sigma^2}{\sqrt{t\sigma^2 (t-k)\,\sigma^2}} = \frac{\sqrt{(t-k)}}{\sqrt{t}}$$

  - Autocorrelations die out slowly even for non-stationary series
  - Difficult to distinguish a true unit root process from a process with a near unit root
  - Examining the ACF is therefore only a first indication

- **Dickey-Fuller test**

How do we test for a unit root?

Figure 12 : ACF of a st a non-stationary series



Basic Dickey-Fuller Test

Consider an AR(1) process

$$y_t = \alpha_1 y_{t-1} + \varepsilon_t$$

where ε_t is a white noise error term.

It seems obvious to use the OLS estimate $\widehat{\alpha}_1$ for α_1 and the corresponding standard error to test the unit root hypothesis

$$H_0: \quad \alpha_1 = 1$$
$$H_1: \quad \alpha_1 < 1$$

by calculating a **standard t-statistic** given by

$$\tau = \frac{\widehat{\alpha}_1 - 1}{\mathrm{se}\left(\widehat{\alpha}_1\right)}$$

Basic Dickey-Fuller Test

We know that

- $\widehat{\alpha}_1$ is consistent irres[...] true value of α_1

But

- Under $H_0 : \alpha_1 = 1$ the standard t-ratio does not have a t-distribution, not even asymptotically.

  The reason for this is that under the null hypothesis the series is non-stationary, which invalidates standard results on the distribution of the OLS estimator.

So we can use the standard t-statistic, but critical values have to be taken from the appropriate distribution. Dickey and Fuller (1979, 1981) derive the appropriate distribution using Monte Carlo simulation. This distribution is **skewed to the right** so that critical values are smaller than those from the normal distribution.

Basic Dickey-Fuller Test: Alternative τ statistic

Usually, a slightly more [illegible] regression is used, i.e. the AR(1) model is rewritten [illegible]

$$y_t - y_{t-1} = (\alpha_1 - 1)\, y_{t-1} + \varepsilon_t$$

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

with $\gamma = \alpha_1 - 1$. The unit root hypothesis now corresponds to

$$H_0: \quad \gamma = 0$$

$$H_1: \quad \gamma < 0$$

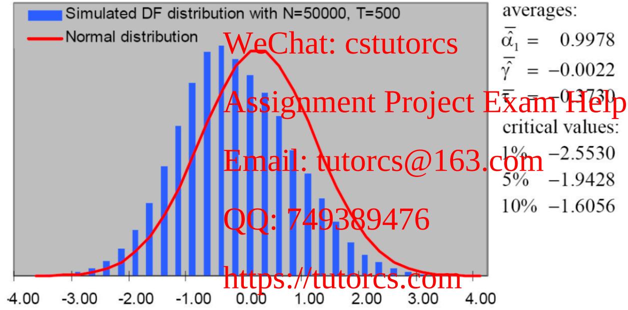
which can be tested by calculating the τ-statistic

$$\tau = \frac{\widehat{\gamma}}{\mathrm{se}(\widehat{\gamma})}$$

which is exactly the same as the one presented above.

Basic Dickey-Fuller Test: Distribution

Figure 13 : Distribution of … in the model: $\Delta y_t = \gamma y_{t-1} + \varepsilon_t$



Basic Dickey-Fuller Test: Example

Example of DF test: [illegible] GDP

ADF Test Statistic	12.72681	1%	[illegible]	[illegible]799
		5%	Critical Value	-1.9421
		10%	Critical Value	-1.6169

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP)
Method: Least Squares
Date: 09/12/07 Time: 11:08
Sample: 1972:1 2007:4
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	0.005522	0.000434	12.72681	0.0000

R-squared	0.011496	Mean dependent var	[illegible]
Adjusted R-squared	0.011496	S.D. dependent var	1.162932
S.E. of regression	1.156228	Akaike info criterion	3.135124
Sum squared resid	191.1715	Schwarz criterion	3.155747
Log likelihood	-224.7289	Durbin-Watson stat	0.765585

Dickey-Fuller Test: Intercept

Extending the DF te[...] an intercept

Now consider the foll[...]rocess

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t$$
$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \varepsilon_t$$

with $\gamma = \alpha_1 - 1$ and ε_t a white noise error term.

- The unit root hypothesis

$$H_0: \quad \gamma = 0$$
$$H_1: \quad \gamma < 0$$

can be tested by calculating the τ_μ-statistic

$$\tau_\mu = \frac{\widehat{\gamma}}{\text{se}\left(\widehat{\gamma}\right)}$$

Dickey-Fuller Test with Intercept: Example

Example of DF test: [illegible] GDP

ADF Test Statistic	1.685221	[illegible]	-3.4767
		5% Critical Value	-2.8815
		10% Critical Value	-2.5773

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP)
Method: Least Squares
Date: 09/12/07 Time: 11:09
Sample: 1972:1 2007:4
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	0.003392	0.002013	1.685221	0.0941
C	0.484439	0.446942	1.083895	0.2802

R-squared	0.019608	Mean dependent var	[illegible]
Adjusted R-squared	0.012703	S.D. dependent var	1.162932
S.E. of regression	1.155522	Akaike info criterion	3.140773
Sum squared resid	189.6029	Schwarz criterion	3.182021
Log likelihood	-224.1357	F-statistic	[illegible]
Durbin-Watson stat	0.770242	Prob(F-statistic)	0.094142

Dickey-Fuller Test: Intercept and Trend

Extending the DF te[illegible] an intercept and trend

Now consider the foll[illegible] process

$$y_t = \alpha_0 + \phi t + \alpha_1 y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \alpha_0 + \phi t + \gamma y_{t-1} + \varepsilon_t$$

with $\gamma = \alpha_1 - 1$ and ε_t a white noise error term.

- The unit root hypothesis

$$H_0: \quad \gamma = 0$$

$$H_1: \quad \gamma < 0$$

can be tested by calculating the τ_τ-statistic

$$\tau_\tau = \frac{\widehat{\gamma}}{\text{se}\left(\widehat{\gamma}\right)}$$

Dickey-Fuller Test with Intercept: Example

Example of DF test: [illegible] GDP

ADF Test Statistic	-0.251139	1% Cr[illegible]	[illegible]41
		5% Cr[illegible]	[illegible]15
		10% Critical Value	-3.1451

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP)
Method: Least Squares
Date: 09/12/07 Time: 11:09
Sample: 1972:1 2007:4
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.004263	0.016976	-0.251139	0.8021
C	1.509957	2.302070	0.655912	0.5129
@TREND(1972:1)	0.008874	0.019538	0.454167	0.6504

R-squared	0.021040	Mean dependent var	1.219947
Adjusted R-squared	0.007154	S.D. dependent var	1.162932
S.E. of regression	1.158765	Akaike info criterion	3.153200
Sum squared resid	189.3259	Schwarz criterion	3.215071
Log likelihood	-224.0304	F-statistic	1.515162
Durbin-Watson stat	0.765485	Prob(F-statistic)	0.223324

Augmented DF Test

The Augmented Dickey-Fuller test

Now consider an AR(p) process

$$y_t = \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + \varepsilon_t$$

where ε_t is a white noise error term.

This series has at least one unit root if the autoregressive parameters sum up to one, i.e. $\sum_{i=1}^{p} \alpha_i = 1$.

Therefore, the unit root hypothesis corresponds to

$$H_0 : \quad \sum_{i=1}^{p} \alpha_i = 1$$

$$H_1 : \quad \sum_{i=1}^{p} \alpha_i < 1$$

Augmented DF Test

Thus, the AR(p) model can be rewritten as

$$\Delta y_t = \gamma y_{t-1} + \varphi_1 \Delta y_{t-1} + \varphi_2 \Delta y_{t-2} + \ldots + \varphi_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

with $\gamma = \left(\sum_{j=1}^{p} \alpha_j\right) - 1$

$$\varphi_i = -\sum_{j=i+1}^{p} \alpha_j$$

The unit root hypothesis again corresponds to

$$H_0: \quad \gamma = 0$$
$$H_1: \quad \gamma < 0$$

and can be tested by calculating the t-statistic

$$\tau = \frac{\widehat{\gamma}}{\mathrm{se}\left(\widehat{\gamma}\right)}$$

Augmented Dickey-Fuller Test: Example

Example of ADF(1) test on Australian GDP

ADF Test Statistic	-2.020490	1% Critical Value*	[illegible]0241
		5% Critical Value	[illegible]4415
		10% Critical Value	-3.1451

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP)
Method: Least Squares
Date: 09/12/07 Time: 11:06
Sample: 1972:1 2007:4
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.027162	0.013443	-2.020490	0.0462
D(GDP(-1))	0.639385	0.066608	9.599179	0.0000
C	3.994947	1.812691	2.203877	0.0292
@TREND(1972:1)	0.032667	0.015427	2.117464	0.0360

R-squared	0.409615	Mean dependent var	1.219947
Adjusted R-squared	0.396964	S.D. dependent var	1.162932
S.E. of regression	0.903080	Akaike info criterion	2.661373
Sum squared resid	114.1774	Schwarz criterion	2.743868
Log likelihood	-187.6188	F-statistic	32.37782
Durbin-Watson stat	1.847073	Prob(F-statistic)	0.000000

Augmented Dickey-Fuller Test: Example

Example of ADF(3) t[illegible]lian GDP

ADF Test Statistic	-1.015628	1% Critical V[illegible]	[illegible]
		5% Critical V[illegible]	[illegible]
		10% Critical V[illegible]	[illegible]

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP)
Method: Least Squares
Date: 09/12/07 Time: 11:07
Sample: 1972:1 2007:4
Included observations: 144

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.014101	0.013884	-1.015628	0.3116
D(GDP(-1))	0.676544	0.082297	8.220781	0.0000
D(GDP(-2))	0.022566	0.100464	0.224619	0.8226
D(GDP(-3))	-0.221547	0.084397	-2.625060	0.0096
C	2.387817	1.855846	1.286647	0.2004
@TREND(1972:1)	0.018272	0.015869	1.151403	0.2516

R-squared	0.445951	Mean dependent var	1.219947
Adjusted R-squared	0.425876	S.D. dependent var	1.162932
S.E. of regression	0.881165	Akaike info criterion	2.625630
Sum squared resid	107.1503	Schwarz criterion	2.749372
Log likelihood	-183.0453	F-statistic	22.21506
Durbin-Watson stat	2.082005	Prob(F-statistic)	0.000000

Model Selection for ADF

Selecting the appropriate model

The appropriate specification of the ADF test and the appropriate critical values depend on the DGP of the process under consideration.

In practice, we usually do not know the true DGP and therefore need to select the appropriate

- order of the AR process
- number of deterministic terms

Model Selection for ADF

Selecting the appropr[illegible]ministic terms

In general, we don't kr[illegible]n intercept and/or a linear trend should be included in the model.

Again we face a **trade-off**:

- On the one hand, it seems reasonable to estimate the most general model:

$$\Delta y_t = \alpha_0 + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \varepsilon_t$$

The problem with this approach is that it may reduce the power of the DF test, as including additional deterministic regressors:

- Reduces the degrees of freedom
- Increases the critical values for testing $\gamma = 0$, resulting in (dramatically) wider confidence intervals

Model Selection for ADF

- On the other hand, [illegible]ely omitting deterministic terms also results i[illegible] of the power of the test
  - If the estimated [illegible]ppropriately omits an intercept term, the ADF-test is consistent but its finite sample power is adversely affected and decreases as the magnitude of the coefficient on the omitted intercept increases.
  - If the estimated regression inappropriately omits a trend term, the power of the ADF-test goes to zero as the sample size increases

The implication of this trade-off is that possibly the null hypothesis of a unit root may not be rejected due to misspecification of the deterministic part of the model, as both too few and too many deterministic regressors reduce the power of the unit root test.

Therefore, it is very important to use a regression that mimics the true DGP as close as possible!

A note on the power of ADF tests?

Power

ADF unit root tests m[illegible]ower against (trend-)stationary alternatives with a root close to unity as in finite samples

- Any near unit root process (i.e. stationary process but with strong inertia) can be (arbitrarily well) approximated by a random walk process
- Any trend-stationary process with a root close to unity can be (arbitrarily well) approximated by a random walk with drift

As a result, **ADF tests too often indicate that a series contains a unit root**.
This is due to the fact that ADF tests have a unit root as the null hypothesis, which implies that we need considerable evidence against the unit root hypothesis in order to be able to reject it.

A note on the power of ADF tests:

Note that although near unit root and true unit root processes on the one hand and trend-stationary and random walk with drift processes on the other hand may be observationally equivalent in finite samples the implications for the properties and for forecasting are very different:

- A unit root process does not converge to a stable mean or deterministic trend while a (trend-)stationary process does