13. Time Series Analysis: Asymptotics
Weakly Dependent and Random Walk Process

Read Wooldridge (2013), Chapter 11, 18.2

Outline: Further Issues in Using OLS with Time Series Data

I. Stationary and Weakly Dependent Time Series
II. Asymptotic Properties of OLS
III. Highly Persistent Time Series
IV. Formal Tests for Unit Root
V. Correcting for Unit Root
VI. Dynamically Complete Model

I. Stationary and Weakly Dependent Time Series

What have we learned?

1) Assume TS.1-TS.6, OLS has exactly same desirable properties that we derived for cross-sectional data.
2) If the errors are **not** drawn from a normal distribution, then we must rely on the central limit theorem – assume large sample properties.
3) In time series analysis, observations tend to correlate across time – TS.5.
4) **Strict exogeneity** (TS.3) is very restrictive. It might not be applicable to static and distributed lag models.

Strict Exogeneity

- Strict exogeneity (TS.3) may not make sense. Consider the static model

  $rgdp_t = \alpha_0 + \alpha_1 ms_t + u$

- Strict exogeneity (TS.3) assuming that u_t is uncorrelated with regressors (X) for all time periods, which implies that

  $E(u_t | ..ms_{t-1}, ms_t, ms_{t+1}, ...) = 0$

  $E(rgpd_t | .. ms_{t-1}, ms_t, ms_{t+1}, ...) = \alpha_0 + \delta_0 ms_t$

- TS.3 rules out the correlation between u_t and ms_{t+1}. Thus, the **dependence of $rgdp_t$ on ms_{t+1} is not allowed**. This may not make sense in time series analysis.

Relax Exogeneity

- Relax: assume TS.3′ that x_t and y_t are weakly dependent.
  With large sample size, assume that u_t is uncorrelated with regressors (x_t) in the same time period only.

4 Concepts

- Goal: to have a general understanding of these concepts:
  1. stationary time series
  2. nonstationary time series
  3. covariance stationary process
  4. weakly dependent time series

1. Stationary Time series

- It refers to the stochastic process whose joint distributions are stable over time.
- Example $x_t = rgdp_t$ $\{x_t: t = 1,2,...\}$
  - The distribution x_t has the same distribution as x_1.
  - The joint distribution of (x_1, x_2) must be the same as the joint distribution of (x_5, x_6) and must be the same as the joint distribution of (x_t, x_{t+1}).
- Thus, stationary implies that the x_t's are **identically distributed** and that the nature of any **correlation between adjacent terms** is the same across all periods.

2. Nonstationary Time Series

- A stochastic process that is not stationary is said to be a nonstationary process
- Violation: a process with a time trend (trending series). It's not stationary. Its mean changes over time.

3. Covariance Stationary Process

- It refers to a stochastic process has a finite second moment [$E(x^2)<\infty$]. This is a weaker form of stationary.
- A stochastic process is covariance stationary if
  i) $E(x_t)$ is constant
  ii) $VAR(x_t)$ is constant
  iii) $Cov(x_t, x_{t+h})$ depends on h, not on t (for all t, h $\geq$0).
- Thus, this weaker form of stationarity requires only that the mean and variance are constant across time, and the covariance just depends on the distance across time

4. Weakly Dependent Time Series

- A stationary time series process is weakly dependent if x_t and x_{t+h} are "almost independent." as h increases without bound.
- Stationary versus Weakly Dependence Time Series
  – *Stationary* – joint distribution of a process
  – *Weakly dependence* - places restrictions on how strongly related x_t and x_{t+h} can be as "h" gets large.

Weakly Dependent Time Series

- A covariance stationary process is weakly dependent if

$$Corr(x_t, x_{t+h}) \rightarrow 0 \text{ as } h \rightarrow \infty.$$

- In other words, covariance stationary sequences are asymptotically uncorrelated.

Example: Moving Average process of Order 1: MA(1)

- $x_t = e_t + \alpha_1 e_{t-1}$; t=1,2, …

  where $\{e_t : t = 0, 1,\ldots\}$ is i.i.d sequence with mean zero and variance σ_e^2. [we write e_t i.i.d $(0, \sigma_e^2)$]

mean:	$E(x_t) = 0$
variance:	$Var(x_t) = (1+\alpha_1^2)\sigma_e^2$
covariance:	$Cov(x_t, x_{t+1}) = \alpha_1\sigma_e^2$
correlation:	$Corr(x_t, x_{t+1}) = \alpha_1/(1+\alpha_1^2)$

- Question: Is x weakly dependent?

Properties: MA(1)

- $x_t = e_t + \alpha_1 e_{t-1}$
- Properties:
  1) x_t is a weighted average of e_t and e_{t-1}
  2) MA(1) process is weakly dependent!
  3) Only adjacent terms are correlated.
  4) After two or more time periods, *x* are independent!

Example: Autoregressive Process of Order 1: AR(1)

- $y_t = \rho y_{t-1} + e_t$

  where $\{e_t\}$ is i.i.d $(0, \sigma_e^2)$ and $|\rho| < 1$ (stability condition)
- Moments:

$E(y_t)$	$=E(y_{t-1})$ when $E(y_t)=0$
$VAR(y_t)$	$= \sigma_y^2 = \sigma_e^2/(1-\rho^2)$
$Cov(y_t, y_{t+h})$	$= \rho^h \sigma_y^2$
$Corr(y_t, y_{t+h})$	$= \rho^h$

- Question: Is *y* weakly dependent?

Properties: AR(1)

- $y_t = \rho y_{t-1} + e_t$
- Properties:
  When $\{y_t\}$ is a stable AR(1) process (i.e., $|\rho|<1$), it is asymptotically uncorrelated or weakly dependent.
  - Let $\rho = 0.9$
  - $corr(y_t, y_{t+1}) = \rho = 0.9$
  - $corr(y_t, y_{t+20}) = \rho^h = 0.122$

Example: Return from SET (Security Exchange of Thailand)

- $return_t = \rho return_{t-1} + e_t$

  where$\{e_t\}$ is i.i.d $(0, \sigma_e^2)$
  $return_t$ – the rate of weekly return from SET index at time t.
  If $|\rho|< 1$, the AR(1) process is *weakly dependent.*
- $\widehat{return}_t = 0.067 return_{t-1}$
  (s.e.) (0.038)
  [t-stat] [1.75]
  [prob.] [.0795]
  n = 689, R^2 = -0.003699, R^2-bar = .0020

Trends Revisited

- A trending series cannot be stationary, since the mean is changing over time.
- A trending series can be weakly dependent.
- If a series is weakly dependent and is stationary about its trend, we will call it a trend-stationary process.
  - As long as a trend (*t*) is included, all is well.

II. Asymptotic Properties of OLS

T.S.1′: Linearity and *weakly dependence*

The stochastic process $\{(x_{t1}, x_{t2},\ldots,x_{tk}, y_t): t = 1, \ldots, n\}$ follows the linear model:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t$$

We add that $\{\underline{\boldsymbol{x}}_t, y_t\}$ is stationary and weakly dependent.

TS.2′: No Perfect Collinearity

No independent variable is constant or a perfect linear combination of the others.

Assumptions: TS.1′- TS.5′

TS.3′: Zero Conditional Mean

$$E(u_t|\underline{x}_t) = 0$$

where $\underline{\boldsymbol{x}}_t = \{x_{t1}, x_{t2},\ldots,x_{tk}\}$ are contemporaneously exogenous.

- Now we do not put restrictions on how u_t is related to $\underline{X}$ in other periods.
  - We relax the strong assumption on *strict exogeneity* of x_{tj}.

OLS Estimators are consistent.

- Theorem 11.1: Consistency of OLS

  If TS.1′-TS.3′ hold, then OLS estimators are consistent.

$$\text{plim}\ \hat{\beta}_j = \beta_j \quad (j=0,1,\ldots,k)$$

Example : Consider the Static Model:

$$inf_t = \beta_0 + \beta_1 ms_t + \beta_2 gdp_t + u_t$$

Notes:

1) $E(u_t | ms_t, gdp_t) = 0$ (weakly dependence)
   u_t are uncorrelated with ms_t and gdp_t (TS.3′)

2) TS.3′ allows the correlation between u_{t-1} and ms_t. *That is,* $corr(u_{t-1}, ms_t) \neq 0$ and $corr(inf_{t-1}, ms_t) \neq 0$.
   This kind of feedback is allowed under TS.3′: Change in ms_t depends on last month's inflation (inf_{t-1}).

- The implies that the TS.3′ allow the feedback from *inf* this period on future values of *ms*.

Example: Consider the AR(1) model

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

Notes:

1) $|\beta_1| < 1$ is the stability condition for weakly dependence.

2) Assume the errors have a zero expected value, given past values of y.
   $E(u_t | y_{t-1}, y_{t-2}, \ldots) = 0$
   $E(y_t | y_{t-1}, y_{t-2}, \ldots) = E(y_t | y_{t-1}) = \beta_0 + \beta_1 y_{t-1}$
   Once y lagged one period has been controlled for, no further lags of y affect the expected value of y_t

3) Strict exogeneity is violated.
   TS.3 requires that, for all t, u_t is uncorrelated with each of (y_0, y_1, …, y_{n-1}). This can not be true.
   Show: $Cov(y_t, u_t) = Var(u_t) > 0$

Assumptions: TS.1′- TS.5′

TS.4′: Homoskedasticity

$VAR(u_t | \underline{x}_t) = \sigma^2$ for all t

The errors are contemporaneously homoskedastic, conditional on regressors only at time t.

TS.5′: No Serial Correlation

$E(u_t u_s | \underline{x}_t, \underline{x}_s) = 0$ for all $t \neq s$

We condition only on regressors only at time t and s.

Asymptotic Normality

- Theorem 11.2: Asymptotic Normality of OLS
  If TS.1′ – TS.5′ hold, OLS estimators are *asymptotically normally distributed.*

- In addition, OLS t, F and LM statistics are valid.

Example: Efficient Markets Hypothesis

- Let y_t be weekly percentage return on NYSE composite index (from Wed close to Wed close)

  $y_t = return_t$

- Efficient Markets Hypothesis

  $E(y_t | y_{t-1}, y_{t-2}, \ldots) = E(y_t)$

  Intuition : past information are useless for prediction.
  If above relationship is false, we could use information on past returns to predict current return.
  (Assume the errors have a zero expected value, given past values of y.)

Example: Efficient Markets Hypothesis

- Specify the AR(1) model

  $return_t = \beta_0 + \beta_1 return_{t-1} + e_t$

  $H_0 : \beta_1 = 0$

  If H_0 is true, then the efficient markets hypothesis is valid.

- $\widehat{return}_t = 0.180 + 0.059 return_{t-1}$
  (s.e.) (0.081) (0.038)
  [t-stat] [2.225] [1.55]
  n = 689, R^2 = .0035, R^2-bar = .0020

- Interpretation:
  1) Explain $\hat{\beta}_1$
  2) t = 1.55 (p-value=.1218). There is a weak (*no*) evidence against the efficient markets (null) hypothesis.

AR(1): Efficient Markets Hypothesis

Dependent Variable: RETURN
Method: Least Squares
Sample(adjusted): 3 691
Included observations: 689 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.179634	0.080742	2.224792	0.0264
RETURN(-1)	0.058898	0.038023	1.549018	0.1218

R-squared	0.003481	Mean dependent var	0.191138
Adjusted R-squared	0.00203	S.D. dependent var	2.11254
S.E. of regression	2.110395	Akaike info criterion	4.334526
Sum squared resid	3059.738	Schwarz criterion	4.347691
Log likelihood	-1491.24	F-statistic	2.399457
Durbin-Watson stat	1.996912	Prob(F-statistic)	0.121838

EM Hypothesis and AR(2)

$return_t = \beta_0 + \beta_1 return_{t-1} + \beta_2 return_{t-2} + e_t$

- If the EM hypothesis is valid under H_0: $\beta_1=0$ and $\beta_2=0$

  $return_t = 0.186 + 0.060 return_{t-1} - 0.038 return_{t-2}$
  (s.e.) (0.081) (0.038) (0.038)
  {t} {2.30} {1.58} {1.00}
  n = 688 R^2 = .0048 R^2bar = 0.038

- Test overall significance: $H_0 : \beta_1 = 0, \beta_2 = 0$
  F-statistic = 1.65; p-value = 0.191.
  Do we reject the efficient markets hypothesis?

AR(2): Efficient Markets Hypothesis

Dependent Variable: RETURN
Method: Least Squares
Sample(adjusted): 4 691
Included observations: 688 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.185748	0.08115	2.288945	0.0224
RETURN(-1)	0.060315	0.038177	1.579865	0.1146
RETURN(-2)	-0.03808	0.038143	-0.998226	0.3185

R-squared	0.004819	Mean dependent var	0.189692
Adjusted R-squared	0.001914	S.D. dependent var	2.113736
S.E. of regression	2.111713	Akaike info criterion	4.337227
Sum squared resid	3054.642	Schwarz criterion	4.356996
Log likelihood	-1489.01	**F-statistic**	**1.658572**
Durbin-Watson stat	1.997717	**Prob(F-statistic)**	**0.191174**

III. Highly persistent time series

- Review: Weakly dependent time series $\{y_t; \underline{x}_t\}$ implies that OLS inference procedures are valid.
  In the case AR(1), the condition is $|\rho| < 1$.
- Terms:
  Highly persistent ≈ strong dependence time series. This series is not weakly dependent.

Highly persistent time series

- Example: AR(1) model with $|\rho| = 1$.

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

  y_t: y at time t
  y_{t-1}: y at previous value
  e_t: i.i.d. $(0, \sigma_e^2)$

- This is a highly persistent time series. This stochastic process $\{y_t\}$ is called a random walk.
- e_t: i.i.d. $(0, \sigma_e^2)$
  This innovations e_t for all t are independently identically distributed with mean zero and constant variance σ_e^2.

A random walk process

- A random walk process is

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

  By repeated substitution, we can write as

  $y_t = e_t + e_{t-1} + \ldots + e_1 + y_0$

- Assume y_0 is nonrandom; mean and variance are

  Mean: $E(y_t) = E(y_0)$
  Variance: $Var(y_t) = t\sigma_e^2$

  This shows that the process can not be stationary!

A random walk displays highly persistent behavior.

- Idea : the value *y* today is highly correlated with the value of *y* in the very distant future.
- *y* that is *h* periods from y_t can be written as

$$y_{t+h} = y_{t+h-1} + e_{t+h}$$
$$y_{t+h} = e_{t+h} + e_{t+h-1} + \dots + e_{t+1} + y_t$$

- for all h ≥ 1

$$E(y_{t+h}) = E(y_t)$$
$$Var(y_{t+h}) = (t+h)\sigma_e^2$$
$$Cov(y_t, y_{t+h}) = t\sigma_e^2$$
$$Corr(y_t, y_{t+h}) = [t/(t+h)]^{1/2}$$

Mean: $E(y_{t+h}) = E(y_t)$.

This means that it does not matter how far in the future we look, the best prediction is today's value *y*.

A random walk is not covariance stationary. Reasons:

1) Variance is not constant. It depends on time t.

$$Var(y_{t+h}) = (t+h)\sigma_e^2$$

2) For any t, h>1, covariance depends, not only on *h* but also on the starting point t.

$$Cov(y_t, y_{t+h}) = t\sigma_e^2$$
$$Corr(y_t, y_{t+h}) = [t/(t+h)]^{1/2}$$

A random walk is not covariance stationary. Reasons:

3) Properties of $corr(y_t, y_{t+h})$

$$\text{corr}(y_t, y_{t+h}) = \sqrt{\frac{t}{t+h}}$$

- For large *t*, corr → 1
- For fixed *t* as h→∞, corr → 0, but it does not do so quickly.
- For large *t*, the more slowly the correlation tends to zero as *h*→∞.
- For large *h*, we can choose a large enough *t* such that corr=1

A unit root process.

- Assuming that $\{e_t\}$ is a weakly dependent series.

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

This is called a unit root process.

- Random walk is a special case of a unit root process

Random Walk vs. Weak Dependence

- Example : GDP time series (Random Walk vs. Weak Dependence)

  Weakly dependence: If GDP next year is **weakly or not** related to GDP thirty years ago. This series is not highly persistent or not strongly dependent.

  Highly persistence: If GDP next year is highly correlated to GDP from many years. This is strongly dependent.

- Key feature of $\{y_t\}$ that has a random walk or unit root process:

  The value of y today is highly correlated with y even in the distant future.

Trending vs. highly persistent

- A trending time series and highly persistent time series are different.

  Examples of Highly persistent time series:

  – inflation rates, interest rates, and unemployment rates.

  Example of trending time series:

  – GDP

- A series can be highly persistent and has a clear trend.

Random Walk with Drift

- $y_t = \alpha_0 + y_{t-1} + e_t$ (*)

  α_0 is the drift term and
  e_t is i.i.d. $(0, \sigma_e^2)$

- By repeated substitution, we write (*) as

  $$y_t = \alpha_0 t + e_t + e_{t-1} + \ldots + e_1 + y_0$$

  Mean: $E(y_t) = \alpha_0 t$ if $y_0 = 0$
  Variance $Var(y_t) = t\sigma_e^2$

Random Walk with Drift

- Mean: $E(y_t) = \alpha_0 t$ implies that
  if $\alpha_0 > 0$ the expected value is growing overtime.
  if $\alpha_0 < 0$ the expected value is shrinking overtime.

- y_{t+h} is y that is h periods from y_t (Random Walk with drift)

  $$y_{t+h} = \alpha_0 h + e_{t+h} + e_{t+h-1} + \ldots + e_{t+1} + y_t$$

  $E(y_{t+h} | y_t) = \alpha_0 h + E(y_t) = \alpha_0 h + y_t$
  $Var(y_{t+h}) = (t+h)\sigma_e^2$
  $Cov(y_t, y_{t+h}) = t\sigma_e^2$
  $Corr(y_t, y_{t+h}) = [t/(t+h)]^{1/2}$

- Interpretation: The best prediction of y_{t+h} is y_t plus the drift $\alpha_0 h$. Variance and Covariance remain the same.

Unit Root Process with intercept

- Unit Root Process with intercept

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

Assume $\{e_t\}$ is any weakly dependent process.

- Here we obtain a whole class of highly persistent time series that also have linearly trending means.

I(0) versus I(1)

- Weakly dependent processes are said to be integrated of order zero, I(0).
- Unit root processes, such as a random walk are said to be integrated of order one, I(1)
- Note that a strong dependence or highly persistent time series violates weakly dependence assumption (TS.1')
  - Thus usual inference procedures (t-test, etc.) are invalid.
  - If we assume TS.1-TS.6 (CLM assumptions), strong dependence in regression analysis poses no problem, but this is unrealistic!

Informal Rules: Differencing or Not

- Deciding whether a time series is I(1): Informal Rules

If $|\rho| < 1$, then the process is I(0)
If $\rho = 1$, then the process is I(1)

- For us,

If $\hat{\rho} > 0.8$, differencing is warranted.

Estimate of ρ

- Find the estimate of ρ (rho)
When AR(1) process is stable,

$$\rho = \text{Corr}(y_t, y_{t-1})$$

This sample correlation $\hat{\rho}$ is called the first order autocorrelation of $\{y\}$.

$$\hat{\rho} = \text{Corr}(y_t, y_{t-1})$$

- Eviews : View/Correlation to find "ρ" - called $\hat{\rho}$

Example: Wages and Productivity

- $\log(hrwage) = \beta_0 + \beta_1 \log(outphr_t) + \beta_2 t + u_t$
  hrwage : average hourly wage
  outphr : output per hour
  t : time trend

- $\widehat{\log(hrwage)} = -5.33 + 1.64 \log(outphr_t) - .018t$
  (s.e) (0.37) (0.09) (.0017)
  {t-stat} {-14.23} {17.65} {-10.42}
  n = 41 R^2 = 0.971 R^2bar= 0.970

- Interpretation:
  1) Test and Interpret $\hat{\beta}_1$
  2) Test and Interpret $\hat{\beta}_2$
  3) Interpret R-Squared.

First Order Autocorrelation

$\log(hrwage) = \beta_0 + \beta_1 \log(outphr_t) + \beta_2 t + u_t$

$\hat{\rho}$ = Corr($lhrwage_t$, $lhrwage_{t-1}$)

	LHRWAGE	LHRWAGE_1
LHRWAGE	1	0.993683
LHRWAGE_1	0.993683	1

$\hat{\rho}$ = Corr($loutphr_t$, $loutphr_{t-1}$)

	LOUTPHR	LOUTPHR_1
LOUTPHR	1	0.997966
LOUTPHR_1	0.997966	1

Regress *lhrwage* on *loutphr* and *t*

Dependent Variable: LHRWAGE
Method: Least Squares
Sample: 1 41
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.32845	0.37445	-14.23008	0
LOUTPHR	1.639639	0.093347	17.56492	0
T	-0.01823	0.001748	-10.42776	0

R-squared	0.971222	Mean dependent var	1.493673
Adjusted R-squared	0.969707	S.D. dependent var	0.163977
S.E. of regression	0.02854	Akaike info criterion	-4.20467
Sum squared resid	0.030952	Schwarz criterion	-4.07929
Log likelihood	89.19582	F-statistic	641.2216
Durbin-Watson stat	0.566845	Prob(F-statistic)	0

Example

Unrealistic: $\widehat{hrwage} = -5.33 + 1.64\ loutphr_t + .081\ t$

Find $\hat{\rho}$: 1st order autocorrelation
$\hat{\rho}$ = 0.967 for detrended log(*hrwage*)
$\hat{\rho}$ = 0.945 for detrended log(*outphr*)
(differencing is warranted)

Dependent Variable: LHRWAGE
Method: Least Squares
Date: 09/15/12 Time: 05:11
Sample: 1 41
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.246799	0.027066	46.06557	0.0000
T	0.011756	0.001123	10.46948	0.0000

R-squared	0.737568	Mean dependent var	1.493673
Adjusted R-squared	0.730839	S.D. dependent var	0.163977
S.E. of regression	0.085072	Akaike info criterion	-2.043082
Sum squared resid	0.282254	Schwarz criterion	-1.959493
Log likelihood	43.88317	F-statistic	109.6100
Durbin-Watson stat	0.060849	Prob(F-statistic)	0.000000

Dependent Variable: LOUTPHR
Method: Least Squares
Date: 09/10/11 Time: 18:33
Sample: 1 41
Included observations: 41

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.010183	0.015576	257.4628	0.0000
T	0.018288	0.000646	28.30134	0.0000

R-squared	0.953570	Mean dependent var	4.394233
Adjusted R-squared	0.952379	S.D. dependent var	0.224346
S.E. of regression	0.048957	Akaike info criterion	-3.148185
Sum squared resid	0.093476	Schwarz criterion	-3.064597
Log likelihood	66.53780	F-statistic	800.9657
Durbin-Watson stat	0.102272	Prob(F-statistic)	0.000000

Example

In the "group" window, View/Correlation

$\hat{\rho}$= Corr(*lhrwage_detrend*$_t$, *lhrwage_detrend*$_{t-1}$)

	LHRWAGE_DETREND	LHRWAGE_DETREND_1
LHRWAGE_DETREND	1	0.967159
LHRWAGE_DETREND_1	0.967159	1

$\hat{\rho}$= Corr(*loutphr_detrend*$_t$, *loutphr_detrend*$_{t-1}$)

	LOUTPHR_DETREND	LOUTPHR_DETREND_1
LOUTPHR_DETREND	1	0.945293
LOUTPHR_DETREND_1	0.945293	1

Steps in finding $\hat{\rho}$ for detrended log(outphr)

Step 1: Regress log(*outphr*) on time trend.
Eviews: *loutphr c t*

Step 2: Obtain residuals from regression in step 1. Choose Proc/Make Residuals and name "*resid01*" as *loutphr_detrend*

Step 3: Generate a new series in Group.
Eviews: *loutphr_detrend_1 = loutphr_detrend(-1)*

Step 4: Find sample correlation between *loutphr_detrend* and *loutphr_detrend_1*.
In the "group" window, View/ Correlation

IV. Formal Tests for Unit Root

1. Testing for unit root with no time trend
2. Testing for unit root with linear time trend

1. Testing for unit root with no time trend

- We know that the presence of a unit root implies that a shock toady has a long-lasting impact. Terms:
  - highly persistent time series
  - strong dependent series
  - I(1) series

- Informal Rule to decide whether a series is unit root or I(1), if the first order autocorrelation
  $\hat{\rho} > 0.8$, differencing is warranted

Testing with unit root with no time trend

- Consider the AR(1) model

  (*) $y_t = \alpha + \rho y_{t-1} + e_t$

  where y_0 is the observed initial value. We assume that $\{e_t\}$ denote a process that has zero mean, given past observed y:

  $E(e_t | y_{t-1}, y_{t-2}, \ldots, y_0) = 0$ (18.18)

- Given (18.18), $\{e_t\}$ is said to be a **martingale difference sequence** with respect to $\{y_{t-1}, y_{t-2}, \ldots\}$. If $\{e_t\}$ is assumed to be i.i.d. sequence with zero mean and is independent of y_0, then it also satisfies (18.18).

Unit Root: $\rho=1$.

$y_t = \alpha + \rho y_{t-1} + e_t$

- It has a unit root if and only if $\rho=1$.
  Hypothesis Testing
  H_0: $\rho = 1$; $\{y_t\}$ has a unit root or $\{y_t\}$ is I(1).
  H_1: $\rho < 1$; $\{y_t\}$ is I(0) or weakly dependent.

- Notes
  1) In practice, $0<\rho<1$
  2) $\rho <1$. Usually no problem with a unit root.
  3) $\rho >1$ This case is not considered, as $\{y_t\}$ is explosive.

Let $\theta = \rho - 1$

- A convenient equation for carrying out the unit root test is

  $\Delta y_t = \alpha + \theta y_{t-1} + e_t$

- Hypothesis Testing
  - H_0: $\theta = 0$ $\{y_t\}$ is I(1)
  - H_1: $\theta < 0$ $\{y_t\}$ is I(0)

- The problem is that, under H_0, y_{t-1} is I(1). Thus, the t statistic does not have approximate standard normal distribution even in large sample.

Dickey-Fuller (DF) Test

- The asymptotic distribution of the t statistic under null hypothesis is calculated by Dickey-Fuller (1979) – known as Dickey-Fuller distribution. The asymptotic distribution of the t statistic under H_0 has come to be known as the Dickey-Fuller (DF) test for a unit root.

- Rejection rule:
  - $t_{\hat{\theta}} < c$

  where c = -2.86 is the critical value at the 5 % level in Table 18.2. If c=2 is used, we would reject H_0 **more than 5%** of the time when H_0 is true.

Table 18.2 Asymptotic Critical Values for Unit Root t Test: No Time Trend

	level	1%	2.5%	5%	10%
DF Test- No time trend	**c**	**-3.43**	**-3.12**	**-2.86**	**-2.57**
t-test	C	-2.326	-1.96	-1.645	-1.282

Example: Wages and Productivity
Test unit root no time trend

$\log(hrwage) = \beta_0 + \beta_1 \log(outphr_t) + \beta_2 t + u_t$

hrwage : average hourly wage
outphr : output per hour
t : time trend

Tests for a unit root (no time trend):

1) $\Delta lhrwage_t = \alpha + \theta lhrwage_{t-1} + e_t$

2) $\Delta loutphr_t = \alpha + \theta loutphr_{t-1} + e_t$

Example: Test for a unit root: log(*hrwage*)

- $\widehat{\Delta lhrwage}_t = 0.117915 - 0.07138 lhrwage_{t-1}$
  {t-stat} {4.619} {-4.19171)

- Notes
  1. The coefficient on *lhrwage* shows that the estimate of ρ is
     $\hat{\rho} = 1 + \theta = 1 - .07138 = 0.92862$.
  2. Is it statistically less than one? *What can you say about unit root?*
     c = -3.43 *(α=1%)*
     $t_{\hat{\theta}}$ < c (-4.19 < -3.43)
     We reject H_0: θ =0 in favor of H_1 : θ < 0 at the 1% significance level.

	level	1%	2.5%	5%	10%
DF Test- No time trend	c	-3.43	-3.12	-2.86	-2.57

  3. What is p-value for the coeffecient on $lhrwage_{t-1}$?
     p-value = 0.0021 (use Eviews)

Example: Easy way in Eviews

- $\Delta lhrwage_t = \alpha + \theta lhrwage_{t-1} + e_t$
- Steps
  1. Go to "Workfile …" Open the variable of interest. (i.e., open *lhrwage*)
  2. Choose View/Unit Root Test.
  3. In the "Unit Root Test" Box

     Test type: Augmented Dickey-Fuller
     Test for unit root in: level
     Include in test equation: intercept
     lag length, User specified 0

Test *lhrwage* for a unit root. Choose View/Unit Root Test

Null Hypothesis: LHRWAGE has a unit root
Exogenous: Constant
Lag Length: 0 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.191701	0.0021
Test critical values:	1% level	-3.605593	
	5% level	-2.936942	
	10% level	-2.606857	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LHRWAGE)
Included observations: 40 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LHRWAGE(-1)	-0.071318	0.017014	-4.191701	0.0002
C	0.117915	0.025528	4.619117	0
R-squared	0.316182	Mean dependent var		0.011547

Example: continue

- $\Delta \widehat{loutphr}_t = \alpha + \theta loutphr_{t-1} + e_t$
  $\Delta \widehat{loutphr}_t = .16810 - 0.03404 loutphr_{t-1}$
  {t-stat} {3.823} {-3.40}
  {prob.} {0.0168}

- Questions.

1) What is the estimate of $\hat{\rho}$?
2) What can we say about unit root?

Test *loutphr* for a unit root. Choose View/Unit Root Test

Null Hypothesis: LOUTPHR has a unit root			
Exogenous: Constant			
Lag Length: 0 (Fixed)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.400018	0.0168
Test critical values:	1% level	-3.605593	
	5% level	-2.936942	
	10% level	-2.606857	
*MacKinnon (1996) one-sided p-values.			

Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LOUTPHR)				
Included observations: 40 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOUTPHR(-1)	-0.034037	0.010011	-3.400018	0.0016
C	0.168105	0.04397	3.823181	0.0005
R-squared	0.233255	Mean dependent var		0.018793

2. Test unit root with time trend (t)

- When $\{y_t\}$ has a clear trend
- A trending-stationary process has a linear trend but is I(0) *about its trend*. In running a DF regression, it can be mistaken for a unit root process.
- Thus, we should control for a time trend in DF.

$$\Delta y_t = \alpha + \delta t + \theta y_{t-1} + e_t$$

- H_0: $\theta = 0$, y_t has a unit root (H_0 is rejected). Thus, $\Delta y_t = \alpha + \delta t + e_t$. This implies that the change in y has a mean *linear in t*.
- H_1: $\theta < 0$, $\{y_t\}$ is a trend-stationary process.

Test unit root with t

- $\Delta y_t = \alpha + \delta t + \theta y_{t-1} + e_t$

  H_0: $\theta = 0$
  H_1: $\theta < 0$

- Rejection rule:
  $t_{\hat{\theta}} < c$ where c = -3.41 is the critical value at the 5 % significance level in Table 18.3.

Table 18.3 Asymptotic Critical Values for Unit Root t Test: Linear Time Trend

	level	1%	2.5%	5%	10%
DF Test- No time trend	c	-3.43	-3.12	-2.86	-2.57
DF Test – with time trend	**c**	**-3.96**	**-3.66**	**-3.41**	**-3.12**
t-test	C	-2.326	-1.96	-1.645	-1.282

Example: Wages and Productivity
Test unit root with time trend (δt)

$\log(hrwage) = \beta_0 + \beta_1 \log(outphr_t) + \beta_2 t + u_t$

hrwage : average hourly wage
outphr : output per hour
t : time trend

Tests for a unit root (no time trend):

1) $\Delta lhrwage_t = \alpha + \delta t + \theta lhrwage_{t-1} + e_t$

2) $\Delta loutphr_t = \alpha + \delta t + \theta loutphr_{t-1} + e_t$

Example: augmented DW Test with linear time trend

- $\Delta \log(\widehat{hrwage}_t) = .0535 - 0.0009t - 0.0157\log(hrwage_{t-1})$
  (s.e.) (.0418) (.00047) (.0334)
  {t} {1.28} {-1.91} {-.470}
  [prob] [.208] [.0639] [.641]
  n=40 R^2 = .37758
- Analysis:
  1) What is the estimate of ρ?
  2) Is there any evidence of unit root at the 10% level?

	level	1%	2.5%	5%	10%
DF Test – with time trend	c	-3.96	-3.66	-3.41	-3.12

3) Interpret coefficient on time trend (Note that t-statistic is invalid).

Easy way in Eviews

- $\Delta lhrwage_t = \alpha + \delta t + \theta hrwage_{t-1} + e_t$
- Steps
  1. Open series "*lwagehr*".
  2. Choose View/Unit Root Test...
  3. In the "Unit Root Test" Box

Test type	Augmented Dickey-Fuller
Test for unit root in	Level
Included in test equation:	trend and intercept
lag length, User specified	0

$\Delta lhrwage_t = \alpha + \delta t + \theta hrwage_{t-1} + e_t$

Null Hypothesis: LHRWAGE has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-0.469976	0.981
Test critical values:	1% level	-4.205004	
	5% level	-3.526609	
	10% level	-3.194611	

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

Sample(adjusted): 2 41
Included observations: 40 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LHRWAGE(-1)	-0.01571	0.033433	-0.469976	0.6411
C	0.053513	0.041781	1.280808	0.2082
t	-0.0009	0.000473	-1.910429	0.0639
R-squared	0.37758	Mean dependent var		0.011547

$\Delta \widehat{loutphr}_t = \alpha + \delta t + \theta loutphr_{t-1} + e_t$

Null Hypothesis: LOUTPHR has a unit root			
Exogenous: Constant, Linear Trend			
Lag Length: 0 (Fixed)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.335314	0.8643
Test critical values:	1% level	-4.205004	
	5% level	-3.526609	
	10% level	-3.194611	
*MacKinnon (1996) one-sided p-values.			

Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LOUTPHR)				
Included observations: 40 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOUTPHR(-1)	-0.06289	0.047099	-1.335314	0.1899
C	0.283183	0.188754	1.500278	0.142
t	0.000561	0.000895	0.627213	0.5344
R-squared	0.241319	Mean dependent var		0.018793

Which model do you prefer?

- **Model I**

$\widehat{\log(hrwage)} = -5.33 + 1.64 \log(outphr) - .018t$

(s.e) (0.37) (0.09) (.0017)

{t-stat} {-14.23} {17.65} {-10.42}

n = 41 R^2 = 0.971 R^2bar= 0.970

log(hrwage) has a unit root! or I(1)

log(outphr) has a unit root! or I(1)

V. Correcting for Unit Root

Remedy: transformation I(0)

- Example: A a random walk process

$y_t = y_{t-1} + e_t$ where $\{e_t\}$ i.i.d. $(0,\sigma^2)$,

$\Delta y_t = y_t - y_{t-1} = e_t$

Δy_t is an i.i.d sequence with with mean zero and constant variance σ^2. The first difference is I(0)

- Example: Assume that e_t is any weakly dependent process.
This is y_t is a unit root process.
The first difference $\Delta y_t = e_t$ is I(0) or weakly dependent.

Remedy: transformation I(0)

- Weakly dependent processes are said to be integrated of order zero I(0).

- Goal: Transform to get a unit root process that is weakly dependent by first difference.

$(\Delta y_t = y_t - y_{t-1})$

Remedy: transformation I(0)

- Consider $\log(y_t)$ that is I(1)

  $\log(y_t) = \log(y_{t-1}) + e_t$

  where $\{e_t\}$ is any weakly dependent process.

- The first difference $\Delta\log(y)$ is weakly dependent or I(0)!

  $\Delta\log(y) = \log(y_t) - \log(y_{t-1}) = e_t$

  Note that $\Delta\log(y) \approx [(y - y(-1))]/y(-1)$

Example

- Example: Efficient Markets hypothesis Model

  $return_t = \beta_0 + return_{t-1} + e_t$

  where $return_t = (price - price(-1))/price(-1) \approx \Delta\log(price)$

- Note that we use $return_t$ instead of $price_t$ in testing efficient markets hypothesis.

Another benefit of differencing: It removes any linear time trend.

- Consider a model

  $y_t = \gamma_0 + \gamma_1 t + v_t$ $\quad E(v_t) = 0$

  $y_{t-1} = \gamma_0 + \gamma_1(t-1) + v_{t-1}$

  $\Delta y_t = \gamma_1 + \Delta v_t$

  $E(\Delta y_t) = \gamma_1$ $\quad E(\Delta v_t) = 0$

- $\therefore E(\Delta y_t) = \gamma_1$ is a constant!
  Differencing removes any linear time trend.

Which model do you prefer?

- **Model I**

  $\widehat{\log(hrwage)} = -5.33 + 1.64 \log(outphr) - .018t$
  (s.e) (0.37) (0.09) (.0017)
  {t-stat} {-14.23} {17.65} {-10.42}
  n = 41 $R^2 = 0.971$ R^2bar= 0.970

- **Model II**

  $\widehat{\Delta\log(hrwage)} = -0.0036 + 0.809\,\Delta\log(outphr)$
  (s.e.) (0.42) (0.173)
  [t-stat] [-.867] [4.67]
  n = 40; $R^2 = 0.364$; R^2-bar= 0.348

Remedy: First differences to generate I(0) processes

- $\Delta\log(\widehat{hrwage}) = -0.0036 + 0.809\,\Delta\log(outphr)$
  (s.e.) (0.42) (0.173)
  [t-stat] [-.867] [4.67]
  n = 40; R^2 = 0.364; R^2-bar= 0.348

- Interpret:
  1) Interpret the coefficient on $\Delta\log(outphr)$
  2) Interpret R-squared.
  Notes: $\Delta\log(hrwage)$ is the growth in average hourly wage.

Eviews Notes: you can write $\Delta\log(hrwage)$ as
lhrwage-lhrwage(-1) or **D(*lhrwage*)**

After differencing:
regress *lhrwage-lwrwage(-1)* on *loutphr-loutphr(-1)*

Dependent Variable: LHRWAGE-LHRWAGE(-1)
Method: Least Squares
Sample(adjusted): 2 41
Included observations: 40 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00366	0.00422	-0.867802	0.391
LOUTPHR-LOUTPHR(-1)	0.809315	0.173454	4.665879	0

R-squared	0.364234	Mean dependent var	0.011547
Adjusted R-squared	0.347503	S.D. dependent var	0.020984
S.E. of regression	0.01695	Akaike info criterion	-5.26834
Sum squared resid	0.010918	Schwarz criterion	-5.1839
Log likelihood	107.3669	F-statistic	21.77043
Durbin-Watson stat	1.52638	Prob(F-statistic)	0.000037

After differencing: Note that D(*z*) = z-z(-1)
regress D(*lhrwage*) on D(*loutphr*)

Dependent Variable: D(LHRWAGE)
Method: Least Squares
Sample(adjusted): 2 41
Included observations: 40 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.00366	0.00422	-0.867802	0.391
D(LOUTPHR)	0.809315	0.173454	4.665879	0

R-squared	0.364234	Mean dependent var	0.011547
Adjusted R-squared	0.347503	S.D. dependent var	0.020984
S.E. of regression	0.01695	Akaike info criterion	-5.26834
Sum squared resid	0.010918	Schwarz criterion	-5.1839
Log likelihood	107.3669	F-statistic	21.77043
Durbin-Watson stat	1.52638	Prob(F-statistic)	0.000037

VI. Dynamically Complete Model

- If the model is dynamically complete, there is no serial correlation.
- A dynamically complete model is a model that is correctly specified.
- Some think that all models should be dynamically complete.
  - Serial correlation in the errors of a model is a sign of misspecification of the model.

Simple Static Regression Model

- $y_t = \beta_0 + \beta_1 z_t + u_t$
- For consistency, we need $E(u_t|z_t) = 0$
- Generally, $\{u_t\}$ is serially correlated. But assume
  $E(u_t|z_t, y_{t-1}, z_{t-1}, y_{t-2}, \dots,) = 0$

  (1)TS.5' holds.
  (2) z_t is contemporaneously exogenous.

Simple Static Regression Model

- $y_t = \beta_0 + \beta_1 z_t + u_t$
- To gain insight, we can write above 2 equations as

  $E(u_t|z_t, y_{t-1}, z_{t-1}, y_{t-2}, \dots,) = 0$
  $E(y_t|z_t, y_{t-1}, z_{t-1}, y_{t-2}, \dots,) = E(y_t|z_t) = \beta_0 + \beta_1 z_t$
- Interpretation: Once z_t has been controlled for, no further lags of either y or z help to explain current y. Model 1 is **dynamically complete.**
  This implies that errors are serial uncorrelated.

AR(1) model

- $y_t = \beta_0 + \beta_1 y_{t-1} + u_t$
- Assume that
  $E(u_t|y_{t-1}, y_{t-2}, \dots,) = 0$
  $E(y_t|y_{t-1}, y_{t-2}, \dots,) = E(y_t|y_{t-1}) = \beta_0 + \beta_1 y_{t-1}$
- Note that this is the same thing as assuming that only one lag of y appears in $E(y_t|y_{t-1}, y_{t-2}, \dots,)$ or there is no serial correlation.
- Interpretation: Once y_{t-1} has been controlled for, no further lags of y affect current y. Model 2 is **dynamically complete!**

General model

- $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + u_t$
  $\underline{\boldsymbol{x}}_t = (x_{t1}, .., x_{tk})$ may contain lags of either z or y affecting y.
- Assume that
  $E(u_t|\underline{\boldsymbol{x}}_t, y_{t-1}, \underline{\boldsymbol{x}}_{t-1}, \dots,) = 0$
  $E(y_t|\underline{\boldsymbol{x}}_t, y_{t-1}, \underline{\boldsymbol{x}}_{t-1}, \dots,) = E(y_t|\underline{\boldsymbol{x}}_t)$
- Interpretation: Further lags of either y or explanatory variables do not help to explain current y. Model 3 is **dynamically complete**!

Recap on TS: Asymptotics

- Stationary and Weakly Dependent Time Series
- Asymptotic Properties of OLS
- Highly Persistent Time Series
- Formal Tests for Unit Root
- Dynamically Complete Model