Nonstationary time series models. Unit root tests

Nuno Sobreira

ISEG - School of Economics & Management

Introduction

- Until now, we either assumed that the time series of interest was stationary (ARMA models) or that a simple transformation (logs, first-differences) was enough to stationarize the time series.
- However, many economic and financial time series are clearly not stationary and so given the importance of this feature we reserve this set of slides to provide a more detailed and structural analysis to this problem.
- Hence, the objective of this group of slides is to:
  - (a) Identify the consequences of ignoring the stationary problem. The most hazardous case – the spurious regression problem – is analyzed in some detail.
  - (b) Present the most typical deviations from the stationary assumption in economic/financial time series and the appropriate transformation.
  - (c) Explain the reasons that make each transformation so effective, in particular, the log or the first-difference transformation.
  - (d) Explain the logic of unit root tests and how this class of statistics can be used to test if a time series is stationary or nonstationary in the mean (unit root/trend stationary).

The spurious regression problem I

- The researcher has to be very cautious with the econometric analysis of **nonstationary time series** .
- If the nonstationarity of the time series is ignored then:
  (a) The estimators of the linear regression model **may be** inconsistent and even divergent. Moreover, the standard (normal) critical values of the usual t, F, LM statistics are not correct anymore.
  (b) The forecast accuracy of the model may be seriously affected.
- We ilustrate item (a) using a simple but somewhat shocking example with real time series data.
- What happens when we run a linear regression model between variables with a **unit root**?

The spurious regression problem II

- Consider two presumably uncorrelated variables:
  - $CONS_t$: Real private consumption
  - $CORVM_t$: number of breeding cormorants

The spurious regression problem III

- Consider the following static regression:

$$\boxed{log(CONS_t) = \beta_0 + \beta_1 log(CORVM_t) + \varepsilon_t} \qquad (1)$$

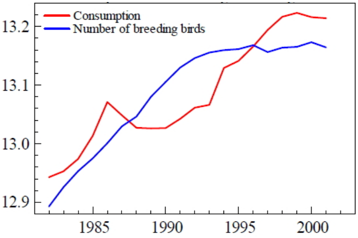
- Naturally $\beta_1 = 0$ and, hopefully, we obtain $R^2 \approx 0$, $\widehat{\beta}_1 \approx 0$ and a very low value of $t_{\widehat{\beta}_1}$.
- With real Danish data, we estimate the static regression in (1) and obtain the following result:

$$\boxed{log(CONS_t) = 12.145 + 0.095 log(CORVM_t) + \widehat{\varepsilon}_t}$$

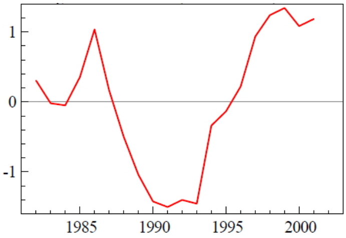
- $t_{\widehat{\beta}_0} = 80.90$, $t_{\widehat{\beta}_1} = 6.30$ and $R^2 = 0.688$!
- The model suggests a clear positive relation between the number of birds and aggregate consumption!
- Furthermore, the number of birds can account for a large proportion of the variation in consumption!

The spurious regression problem IV

- The model seems to be perfectly adequate to explain the Denmark cosumption behaviour but this is obviously a totally absurd result/regression!



(a) Plot of $log(CONS_t)$ and $log(CORVM_t)$



(b) Residuals from the regression of $log(CONS_t)$ on $log(CORVM_t)$

- From visual inspection of the time series plot and unit root tests (not reported here), we conclude that both time series have a unit root.
- The presence of a unit root generates these meaningless results.

The spurious regression problem V

- This set of results is known as the **spurious regression** problem. It occurs when the variables of the model are nonstationary in the mean.
- In general, no matter the relevance of the model, a static regression between unit root variables will result in:
  (a) OLS estimators with high and statistically significant $t_{\widehat{\beta}_j}$, $j = 1, \ldots, k$.
  (b) Very high R^2 and close to 1.
  (c) Very low Durbin-Watson (DW) statistic and close to 0.

Important Warning 1

This is a very important problem but the **inconsistency** of the estimators **only holds** for **static regressions**.

The spurious regression problem VI

Important Warning 2

⇒ **Dynamic models** such as **ARIMA** or **regressions with lagged** dependent or independent **variables** provide **inefficient** but **consistent** estimators even if the **nonstationarity** is **ignored** or **not properly modelled**.

⇒ The **standard statistical inference tools** – **t, F, LM tests** – do **not have** standard **normal distributions** and, in particular, the **p-values** provided by **EViews** are **totally incorrect** which restricts considerably the statistical analysis of the model.

Nonstationarity types I

- The theory underlying the class of ARMA models relies on the assumption that the time series process, X_t, is stationary:

$$E(X_t) = \mu < \infty \text{ , i. e., constant over time}$$

$$Var(X_t) = E\left[(X_t - \mu)^2\right] = \gamma_0 < \infty \text{ , i. e., constant over time}$$

$$Cov(X_t, X_{t-k}) = E\left[(X_t - \mu)(X_{t-k} - \mu)\right] = \gamma_k \text{ , i. e., constant over time}$$

- However, most observed economic and financial time series do not seem to be well characterized as a stationary process which poisons the reliability of the results.
- Fortunately, there are models that can be constructed to describe the many different ways that a time series can be nonstationary.
- These models also show how the nonstationarity problem may be treated in practical applications, in particular, how to transform a nonstationary time series into a stationary one.

Nonstationarity types II

- The time series can be nonstationary in many different ways, and we present the most typical deviations from the stationary assumption in financial/economic time series:
  1. Nonstationarity in variance
  2. Nonstationarity in mean: Linear deterministic trends and stochastic trends
- Identifying the nature of nonstationarity in the time series is an essential step for a correct subsequent data analysis.
- The main challenge is that each form of nonstationarity demands a different transformation.
- The application of an unnecessary or wrong transformation may have detrimental effects in the statistical properties of the econometric model and forecast performance.

Nonstationarity in variance* (Optional) I

- Many time series have nonconstant variance and it is very common for the variance to change as its level, μ_t, changes:

$$\boxed{Var\left(X_t\right) = cf\left(\mu_t\right)} \tag{2}$$

  for some $c > 0$ and a function f.
- We want to find a transformation, T, that renders a new series, $T(X_t)$, with constant variance.
- The form of the transformation can be obtained with a first order Taylor approximation argument. We start by approximating $T(X_t)$ about the point μ_t which gives that:

$$\boxed{T(X_t) \approx T(\mu_t) + T'(\mu_t)\left(X_t - \mu_t\right)} \tag{3}$$

- According to equations (2) and (3), we have that the variance of the transformed series can be approximated by:

$$\boxed{Var\left[T\left(X_t\right)\right] \approx c\left[T'\left(\mu_t\right)\right]^2 f\left(\mu_t\right)}$$

Nonstationarity in variance* (Optional) II

- Thus, to stabilize the variance, the transformation must be choosen such that:

$$T'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}}$$

or, in other words:

$$T(\mu_t) = \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t$$

- For example, if the standard deviation is proportional to the level so that $Var(X_t) = c^2\mu_t^2$, we have:

$$T(\mu_t) = \int \frac{1}{\sqrt{\mu_t^2}} d\mu_t = \ln(\mu_t)$$

Nonstationarity in variance* (Optional) III

- In general, we may use the Box-Cox power transformation:

$$T\left(X_t\right) = \frac{X_t^{\lambda} - 1}{\lambda}$$

- Tipical values of λ are $-1, -0.5, 0.5$ and 1. If $\lambda \to 0$, $T\left(X_t\right) = \ln\left(X_t\right)$
- In fact, the **most used transformation** is the **logarithm transformation**.

Nonstationarity in mean – trend-stationary models I

- Many economic time series are not in accordance with the constant mean assumption as they have a tendency to systematically increase or decrease over time.
- One way to describe this pattern is to have a model where the mean is not constant but instead follows a linear trend or even a polynomial trend:

$$\boxed{X_t = \beta_0 + \beta_1 t + u_t}$$

$$\boxed{X_t = \beta_0 + \beta_1 t + +\beta_2 t + \ldots + \beta_p t^p + u_t}$$

  where u_t is a zero mean stationary process.
- **Quadratic** or **higher order polynomial trends** are very rare in economic/financial time series data and so we focus on models with a **linear deterministic trend**.

Nonstationarity in mean – trend-stationary models II

- To complete the description of the model and to make it useful in practice, we have to be more concrete as regards to the stochastic part of the model, u_t.
- In theory, the most simple case is a model where the stochastic component, u_t, follows a white-noise process:

$$\boxed{X_t = \beta_0 + \beta_1 t + u_t,\ u_t \overset{w.n.}{\sim} (0, \sigma_u^2)} \tag{4}$$

- However, in practical applications, it is almost unrealistic to have a time series in which the deviations from the trend, $X_t - \beta_0 - \beta_1 t$, follow a white noise process. Time dependence is a proeminent phenomenon as we constantly emphasize during the course.

Nonstationarity in mean – trend-stationary models III

- The most common situation is to find some pattern in the ACF/PACF of the detrended variable, $X_t - \beta_0 - \beta_1 t$, that suggests a stationary and invertible ARMA representation for the detrended variable:

$$\begin{cases} X_t = \beta_0 + \beta_1 t + u_t \\ \phi(L)u_t = \theta(L)\varepsilon_t, \ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2) \end{cases} \tag{5}$$

  where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$. Moreover, the roots of the polynomials $\phi(z)$ and $\theta(z)$ are outside the unit circle.
- This model representation may seem awkward but it is possible to show that the process (5) can be written in the usual ARMA representation but with a linear trend in the deterministic component:

$$\boxed{\phi(L)X_t = \alpha_0 + \alpha_1 t + \theta(L)\varepsilon_t}$$

  where $\alpha_0 = \beta_0\phi(1) + \beta_1 \sum_{j=1}^{p} j\phi_j$ and $\alpha_1 = \beta_1\phi(1)$.

Nonstationarity in mean – trend-stationary models IV

- For both models (4) and (5), it should be clear that the mean changes over time $E(X_t) = \beta_0 + \beta_1 t$ since $E(u_t) = 0$.
- The detrended time series, $X_t - E(X_t) = X_t - \beta_0 - \beta_1 t$ is a stationary and invertible process given that the **roots** of the polynomials $\phi(z)$ and $\theta(z)$ are **outside** the unit circle.
- A time series process with the two features described above is called a **trend-stationary process**.

Definition (Trend-stationary processes)

A time series process is **trend-stationary** if:

(a) $E(X_t)$ is a function of time, $E(X_t) = f(t)$ and, in general, $E(X_t) = \beta_0 + \beta_1 t$ with $\beta_0 > 0$ and $\beta_1 > 0$, i. e., the unconditional mean grows linearly over time.

(b) $X_t - E(X_t) \sim$ stationary and invertible *ARMA*

Nonstationarity in mean – trend-stationary models V

- It is immediate that both models (4) and (5) are **trend-stationary processes**.
- Naturally, in practice we don't know neither the true values of β_0 and β_1 (needed to obtain $X_t - \beta_0 - \beta_1 t$) nor the theoretical ACF and PACF of the detrended process.
- Hence, we obtain sample estimates of β_0 and β_1 with the following linear regression model estimated by OLS:

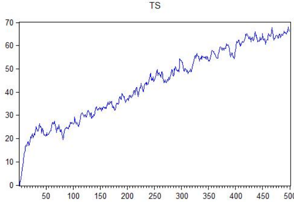
$$\boxed{X_t = \beta_0 + \beta_1 t + u_t}$$

  and proceed with the analysis of the sample ACF/PACF of the residual series $\widehat{u}_t = X_t - \widehat{\beta}_0 - \widehat{\beta}_1 t$.
- We then use the Box Jenkins methodology to choose the orders q and p.

Simulated sample of a trend-stationary process

Graph and correlogram of a trend-stationary process with

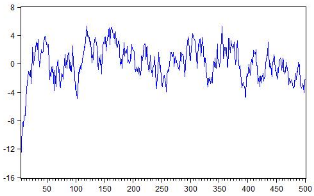
$X_t = 2 + 0.01t + 0.9X_{t-1} + u_t, \; u_t \overset{i.i.d}{\sim} N(0,1)$



Sample: 1 500
Included observations: 500

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.988	0.988	490.98	0.000
		2	0.976	0.003	971.26	0.000
		3	0.964	-0.001	1441.1	0.000
		4	0.953	0.009	1900.8	0.000
		5	0.943	0.034	2351.6	0.000
		6	0.933	0.029	2794.1	0.000
		7	0.924	0.017	3228.9	0.000
		8	0.916	0.037	3656.9	0.000
		9	0.909	0.042	4079.2	0.000
		10	0.902	0.009	4495.9	0.000
		11	0.895	-0.001	4907.0	0.000
		12	0.889	0.030	5313.1	0.000
		13	0.883	0.029	5714.7	0.000
		14	0.877	-0.004	6111.8	0.000
		15	0.871	0.008	6504.4	0.000
		16	0.865	0.012	6892.8	0.000
		17	0.860	0.017	7277.1	0.000
		18	0.854	-0.023	7656.9	0.000
		19	0.848	-0.007	8032.0	0.000
		20	0.842	0.027	8402.9	0.000
		21	0.837	0.023	8770.2	0.000
		22	0.833	0.026	9134.3	0.000
		23	0.828	-0.001	9495.3	0.000
		24	0.824	0.004	9853.0	0.000
		25	0.819	0.003	10208.	0.000
		26	0.814	-0.027	10559.	0.000
		27	0.809	0.024	10906.	0.000
		28	0.805	0.016	11251.	0.000
		29	0.800	-0.012	11592.	0.000
		30	0.797	0.037	11931.	0.000
		31	0.792	-0.019	12267.	0.000
		32	0.788	0.003	12600.	0.000
		33	0.783	-0.023	12929.	0.000
		34	0.777	-0.034	13255.	0.000
		35	0.772	0.014	13577.	0.000
		36	0.767	0.021	13895.	0.000

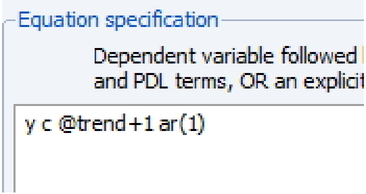
Graph and correlogram of the detrended series $\widehat{u}_t = X_t - \widehat{\beta}_0 - \widehat{\beta}_1 t$

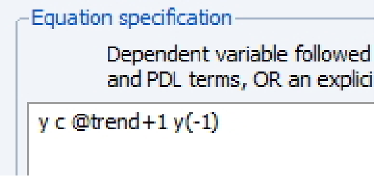


Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.873	0.873	380.68	0.000
		2	0.762	-0.004	670.91	0.000
		3	0.674	0.041	898.65	0.000
		4	0.604	0.030	1081.7	0.000
		5	0.539	-0.005	1227.7	0.000
		6	0.488	0.034	1347.8	0.000
		7	0.431	-0.045	1441.8	0.000
		8	0.367	-0.057	1510.0	0.000
		9	0.322	0.039	1562.7	0.000
		10	0.285	0.002	1603.9	0.000
		11	0.247	-0.020	1634.9	0.000
		12	0.212	-0.003	1658.0	0.000
		13	0.175	-0.038	1673.6	0.000
		14	0.146	0.019	1684.5	0.000
		15	0.108	-0.061	1690.5	0.000
		16	0.073	-0.026	1693.2	0.000
		17	0.056	0.051	1694.8	0.000
		18	0.037	-0.028	1695.5	0.000
		19	0.020	0.002	1695.7	0.000
		20	0.021	0.062	1696.0	0.000
		21	0.023	0.007	1696.2	0.000
		22	0.012	-0.039	1696.3	0.000
		23	-0.005	-0.035	1696.3	0.000
		24	-0.003	0.051	1696.3	0.000
		25	-0.007	-0.019	1696.4	0.000
		26	0.003	0.057	1696.4	0.000
		27	0.010	-0.014	1696.4	0.000
		28	0.019	0.028	1696.6	0.000
		29	0.013	-0.042	1696.7	0.000
		30	0.003	-0.029	1696.7	0.000
		31	0.010	0.047	1696.7	0.000
		32	0.007	-0.037	1696.8	0.000
		33	-0.002	-0.027	1696.8	0.000
		34	-0.010	-0.008	1696.8	0.000
		35	-0.015	-0.002	1697.0	0.000
		36	-0.018	0.018	1697.1	0.000

Exercise

1. Collect data from any time series of your interest. Estimate an AR(1) model with a deterministic trend in the following two different ways:





   (a) Compare the EViews output from the two methods. In particular, show how you obtain the coefficient estimates from the 2nd figure with the coefficient estimates from the 1st figure. Justify your answer mathematically.

2. Repeat the same exercise but now with an AR(2) process making the necessary arrangements on the EViews commands.
3. Obtain the general rule that relates the 2 methods for an ARMA(p,q) process with trend.

Difference-stationary processes I

- Another way of modelling a nonstationary time series is to consider a modified form of the ARMA model with the autoregressive parameters not satisfying the stationary conditions.
- For example, consider the AR(1) process ($\phi_0 = 0$ with no loss of generality):

$$X_t = \phi_1 X_{t-1} + \varepsilon_t,\ \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

- This process is stationary if $|\phi_1| < 1$.
- If $\phi_1 > 1$ ($\phi_1 < -1$ is clearly unrealistic) both $E(X_t)$ and $Var(X_t)$ are increasing functions of time and become infinite as $t \to \infty$. You can recall the expressions for $E(X_t)$ and $Var(X_t)$ on the slides from Chapter 2.
- The same phenomenon occurs with general ARMA processes where the (inverse) roots of the AR polynomial $\phi(z)$ are (outside) inside the unit circle.
- Processes with these features are called **explosive**.

Difference-stationary processes II

- **Explosive** processes are **very rare** in economic and financial time series and we will simply **discard** them.
- A **very interesting** and **much more common** case arises when **one** or **more than one** of the AR polynomial **roots** are **unity** and the **others** are **outside** the unit circle.

Definition (Unit Root)

If $z = 1$ is a root of the AR polynomial, *i.e.*, if $z = 1$ is a solution to the equation:

$$\boxed{\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p = 0}$$

then we say that the time series process X_t has a **unit root**.

- Empirically, it is very frequent to find time series with **one unit root**, $z_1 = 1$, and the others **outside the unit circle**, $|z_2| > 1, \ldots, |z_p| > 1$.

Difference-stationary processes III

- In that case, X_t is **nonstationary** but has a **stationary** and **invertible** ARMA representation after taking **first-differences**. To see this clearly suppose that X_t follows the process:

$$\boxed{\phi^*(L)X_t = \phi_0 + \theta(L)\varepsilon_t} \tag{6}$$

  where $z_1, \ldots, z_p$ are the roots of $\phi^*(z)$ with $z_1 = 1$ and $z_2 > 1, \ldots, z_p > 1$. Moreover assume that the roots of $\theta(z)$ are outside the unit circle.

- Now if we factor the unit root from the polynomial $\phi^*(L)$, equation (6) becomes:

$$\boxed{\phi(L)(1-L)X_t = \phi_0 + \theta(L)\varepsilon_t}$$

- Since $z_2, \ldots, z_p$ are the roots from the polynomial $\phi(z)$ and are outside the unit circle, the first-differenced time series $\Delta X_t = (1-L)X_t$ has a stationary ARMA representation:

$$\boxed{\phi(L)\Delta X_t = \phi_0 + \theta(L)\varepsilon_t}$$

Difference-stationary processes IV

- In theory, nothing prevents to have a process with **d** **unit roots**. In that case, the original series X_t is **nonstationary** but the **dth differenced series**, $\Delta^d X_t$ is **stationary** and has a stationary and invertible ARMA representation:

$$\phi(L)\Delta^d X_t = \phi_0 + \theta(L)\varepsilon_t \tag{7}$$

- These processes are called **Integrated Processes of Order** ***d*** (abbreviatedly, **I(d) processes**).
- The model in (7) is referred to as the **Autoregressive Integrated Moving Average** model of **order** $(\mathbf{p}, \mathbf{d}, \mathbf{q})$ and is denoted as the **ARIMA(p,d,q)** model.
- Example: **ARIMA(1,1,1)**

$$X_t = 1.2X_{t-1} - 0.2X_{t-2} + \varepsilon_t - 0.5\varepsilon_{t-1}, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2) \tag{8}$$

Difference-stationary processes V

- The roots of the AR polynomial are $z_1 = 1$ and $z_2 = 5 > 1$. We conclude that the process (8) follows an **ARIMA(1,1,1)** process because:
  1. X_t is **nonstationary** due to **one** unit root ($x_1 = 1$).
  2. ΔX_t is **stationary** ($d = 1$) and follows an **ARMA(1,1)** process.
- To see point 2 clearly we can factor the AR polynomial in (8).
- Given that its roots are $z_1 = 1$ and $z_2 = 5$ we have that $\lambda_1 = \frac{1}{z_1} = 1$ and $\lambda_2 = \frac{1}{z_2} = 0.2$ and so the factorization of $\phi(L)$ becomes:

$$
\begin{aligned}
&(1 - 1.2L + 0.2L^2)X_t = \varepsilon_t - 0.5\varepsilon_{t-1} \\
\Leftrightarrow &(1 - 0.2L)(1 - L)X_t = \varepsilon_t - 0.5\varepsilon_{t-1} \\
\Leftrightarrow &\Delta X_t = 0.2\Delta X_{t-1} + \varepsilon_t - 0.5\varepsilon_{t-1} \implies \Delta X_t \sim ARMA(1,1)
\end{aligned}
$$

Difference-stationary processes VI

- If $d \geq 1$, we can also say that X_t is **difference-stationary** process.

Definition (Difference-stationary processes)

A process X_t is **difference-stationary** or $I(d)$ for $d \geq 1$, if:

(a) X_t has d unit roots, $d \geq 1$.

(b) $\Delta^d X_t \sim$ stationary and invertible *ARMA*

- Some of the most popular difference-stationary processes are:
  1. **Random Walk** (ARIMA(0,1,0) without constant)

$$X_t = X_{t-1} + \varepsilon_t, \ \varepsilon_t \overset{w.n}{\sim} (0, \sigma_\varepsilon^2) \tag{9}$$

  2. **Random Walk with drift** (ARIMA(0,1,0) with constant)

$$X_t = \beta + X_{t-1} + \varepsilon_t, \ \varepsilon_t \overset{w.n}{\sim} (0, \sigma_\varepsilon^2) \tag{10}$$

Difference-stationary processes VII

- Some time series, specially the ones related to the financial sector, are well described by a random walk or a random walk with drift.
- However, in many circumstances after taking first differences, we find that the sample ACF and PACF of ΔX_t do not have a “white-noise" pattern but of some stationary and invertible ARMA(p,q) process.
- In that case the model that best fits the data is given by:

  3. **ARIMA(p,1,q)**

$$\boxed{\phi(L)\Delta X_t = \phi_0 + \theta(L)\varepsilon_t,\ \varepsilon_t \overset{w.n}{\sim} (0, \sigma_\varepsilon^2)} \tag{11}$$

  where the roots of $\phi(z)$ and $\theta(z)$ are **outside** the unit circle.

- In the following slides, we present more details about the properties of the Random Walk with and without drift.
- The knowledge of these properties is useful for a complete understanding of the unit root tests.

Random Walk I

- A Random Walk is a time series process defined by the following equation:

$$X_t = X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2) \tag{12}$$

- To analyse the properties of the Random Walk, it is useful to write equation (12) as a function of the initial value, X_0 and the shocks:

$$X_t = X_0 + \varepsilon_1 + \ldots + \varepsilon_t = X_0 + \sum_{i=1}^{t} \varepsilon_i$$

Properties of the random walk

- Expected value of X_t:

$$E(X_t) = X_0$$

- Variance of X_t:

$$Var(X_t) = t\sigma_\varepsilon^2$$

Random Walk II

- Autocovariances of X_t:

$$\gamma_{t,k} = Cov(X_t, X_{t-k}) = (t-k)\sigma_\varepsilon^2$$

- Autocorrelations of X_t:

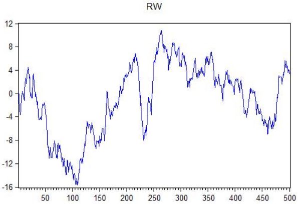
$$\rho_{t,k} = Corr(X_t, X_{t-k}) = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t) \cdot Var(X_{t-k})}} = \sqrt{\frac{t-k}{t}}$$

- Initially, for low values of k, the autocorrelation function is approximately equal to 1.
- As k increases, the autocorrelation function decays very slowly.

Simulated sample of a random walk

Graph and correlogram of a simulated random walk with

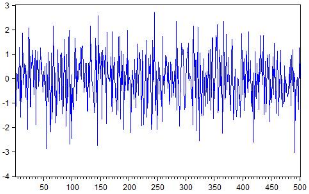
$X_t = X_{t-1} + \varepsilon_t,\ X_0 = 0\ \varepsilon_t \overset{i.i.d}{\sim} N(0,1)$



Sample: 1 500
Included observations: 500

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.986	0.986	488.83	0.000
		2	0.970	-0.077	962.72	0.000
		3	0.952	-0.041	1420.8	0.000
		4	0.934	-0.033	1862.7	0.000
		5	0.916	-0.013	2288.3	0.000
		6	0.898	-0.008	2697.8	0.000
		7	0.880	-0.009	3091.7	0.000
		8	0.863	0.039	3471.2	0.000
		9	0.844	-0.078	3835.2	0.000
		10	0.825	-0.004	4184.0	0.000
		11	0.806	-0.031	4517.3	0.000
		12	0.788	0.035	4836.3	0.000
		13	0.769	-0.031	5141.0	0.000
		14	0.752	0.066	5433.4	0.000
		15	0.734	-0.064	5712.6	0.000
		16	0.716	-0.042	5978.3	0.000
		17	0.700	0.099	6233.0	0.000
		18	0.685	-0.013	6477.1	0.000
		19	0.669	-0.016	6710.5	0.000
		20	0.654	0.025	6934.4	0.000
		21	0.640	0.009	7149.2	0.000
		22	0.627	-0.003	7355.5	0.000
		23	0.615	0.029	7554.3	0.000
		24	0.601	-0.054	7744.6	0.000
		25	0.590	0.076	7928.3	0.000
		26	0.580	0.039	8106.5	0.000
		27	0.571	0.010	8279.5	0.000
		28	0.564	0.049	8448.7	0.000
		29	0.558	0.032	8614.8	0.000
		30	0.554	0.030	8778.5	0.000
		31	0.550	0.008	8940.3	0.000
		32	0.545	-0.056	9099.5	0.000
		33	0.538	-0.072	9254.8	0.000
		34	0.531	0.038	9406.7	0.000
		35	0.526	0.036	9556.0	0.000
		36	0.519	-0.077	9701.5	0.000

Graph and correlogram of the first-differences, $\Delta X_t = X_t - X_{t-1}$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.036	-0.036	0.6537	0.419
		2	-0.053	-0.054	2.0457	0.360
		3	-0.048	-0.052	3.2156	0.360
		4	-0.026	-0.033	3.5659	0.468
		5	-0.034	-0.042	4.1428	0.529
		6	0.043	0.034	5.0843	0.533
		7	0.042	0.039	5.9888	0.541
		8	-0.032	-0.029	6.5045	0.591
		9	-0.003	0.001	6.5077	0.688
		10	0.018	0.020	6.6686	0.756
		11	0.002	0.006	6.6710	0.825
		12	0.036	0.039	7.3495	0.834
		13	-0.027	-0.027	7.7111	0.862
		14	0.035	0.040	8.3432	0.871
		15	-0.011	-0.003	8.4054	0.907
		16	-0.078	-0.079	11.556	0.774
		17	-0.033	-0.037	12.105	0.794
		18	0.021	0.007	12.332	0.830
		19	-0.042	-0.053	13.257	0.825
		20	0.025	0.015	13.581	0.851
		21	0.045	0.032	14.658	0.840
		22	-0.004	0.002	14.667	0.876
		23	-0.070	-0.057	17.237	0.797
		24	0.046	0.038	18.342	0.786
		25	-0.056	-0.055	20.015	0.746
		26	0.047	0.048	21.186	0.732
		27	-0.010	-0.015	21.242	0.775
		28	0.057	0.058	22.948	0.736
		29	0.032	0.049	23.488	0.754
		30	-0.049	-0.040	24.762	0.737
		31	0.006	0.015	24.779	0.778
		32	-0.004	-0.010	24.786	0.815
		33	0.012	0.005	24.858	0.845
		34	-0.007	-0.007	24.885	0.873
		35	0.017	0.004	25.035	0.894
		36	-0.001	-0.003	25.035	0.915

Random walk with drift

- A random walk with drift is a time series process defined by the following equation:

$$X_t = \beta + X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$$

- To study the properties of this process, we write equation (14) as a function of the deterministic term and the random shocks:

$$X_t = X_0 + \beta + \varepsilon_1 + \ldots + \beta + \varepsilon_t = X_0 + \beta t + \sum_{i=1}^{t} \varepsilon_i$$

Properties of the random walk with drift

- Expected value of X_t:

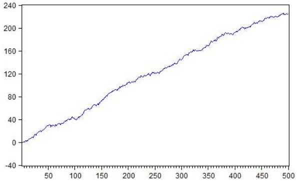
$$E(X_t) = X_0 + \beta t$$

- Variance, autocovariances and autocorrelations equal to the random walk.

Simulated sample of a random walk with drift

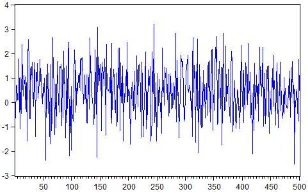
Graph and correlogram of a simulated random walk with drift with

$X_t = 0.5 + X_{t-1} + \varepsilon_t, \; X_0 = 0 \; \varepsilon_t \overset{i.i.d}{\sim} N(0,1)$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.994	0.994	497.40	0.000
		2	0.989	-0.007	990.16	0.000
		3	0.983	0.005	1478.4	0.000
		4	0.978	-0.006	1962.0	0.000
		5	0.972	-0.006	2441.1	0.000
		6	0.966	-0.003	2915.6	0.000
		7	0.961	0.002	3385.6	0.000
		8	0.955	-0.007	3851.0	0.000
		9	0.949	-0.011	4311.8	0.000
		10	0.944	-0.012	4767.9	0.000
		11	0.938	-0.006	5219.3	0.000
		12	0.932	-0.006	5665.9	0.000
		13	0.926	0.003	6107.8	0.000
		14	0.920	0.004	6545.2	0.000
		15	0.914	-0.006	6978.0	0.000
		16	0.909	0.004	7406.2	0.000
		17	0.903	-0.004	7830.0	0.000
		18	0.897	0.004	8249.4	0.000
		19	0.892	-0.008	8664.3	0.000
		20	0.886	0.000	9074.8	0.000
		21	0.880	-0.002	9480.9	0.000
		22	0.875	-0.015	9882.6	0.000
		23	0.869	0.000	10280.	0.000
		24	0.863	0.005	10673.	0.000
		25	0.858	-0.000	11061.	0.000
		26	0.852	-0.006	11445.	0.000
		27	0.846	-0.010	11825.	0.000
		28	0.840	-0.002	12201.	0.000
		29	0.835	0.000	12572.	0.000
		30	0.829	-0.003	12939.	0.000
		31	0.823	0.000	13302.	0.000
		32	0.818	-0.008	13661.	0.000
		33	0.812	-0.006	14015.	0.000
		34	0.806	-0.004	14365.	0.000
		35	0.801	0.008	14711.	0.000
		36	0.795	-0.011	15053.	0.000

Graph and correlogram of the first-differences, $\Delta X_t = X_t - X_{t-1}$



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.036	-0.036	0.6537	0.419
		2	-0.053	-0.054	2.0457	0.360
		3	-0.048	-0.052	3.2156	0.360
		4	-0.026	-0.033	3.5659	0.468
		5	-0.034	-0.042	4.1428	0.529
		6	0.043	0.034	5.0843	0.533
		7	0.042	0.039	5.9888	0.541
		8	-0.032	-0.029	6.5045	0.591
		9	-0.003	0.001	6.5077	0.688
		10	0.018	0.020	6.6686	0.756
		11	0.002	0.006	6.6710	0.825
		12	0.036	0.039	7.3495	0.834
		13	-0.027	-0.027	7.7111	0.862
		14	0.035	0.040	8.3432	0.871
		15	-0.011	-0.003	8.4054	0.907
		16	-0.078	-0.079	11.556	0.774
		17	-0.033	-0.037	12.105	0.794
		18	0.021	0.007	12.332	0.830
		19	-0.042	-0.053	13.257	0.825
		20	0.025	0.015	13.581	0.851
		21	0.045	0.032	14.658	0.840
		22	-0.004	0.002	14.667	0.876
		23	-0.070	-0.057	17.237	0.797
		24	0.046	0.038	18.342	0.786
		25	-0.056	-0.055	20.015	0.746
		26	0.047	0.048	21.186	0.732
		27	-0.010	-0.015	21.242	0.775
		28	0.057	0.058	22.948	0.736
		29	0.032	0.049	23.488	0.754
		30	-0.049	-0.040	24.762	0.737
		31	0.006	0.015	24.779	0.778
		32	-0.004	-0.010	24.786	0.815
		33	0.012	0.005	24.858	0.845
		34	-0.007	-0.007	24.885	0.873
		35	0.017	0.004	25.035	0.894
		36	-0.001	-0.003	25.035	0.915

Trend-stationary versus difference-stationary processes* (Optional)

- In the previous slides we presented two different classes of non stationary time series models: trend-stationary and difference-stationary processes.
- But why is it so **important to analyse** if the series of interest is **trend-stationary** or **difference-stationary**?
- In the next slides, we motivate the importance of this distinction with simple examples.
- Finally, we introduce the **unit root tests** that allow the practitioner to classify a time series as stationary, trend-stationary or difference-stationary with formal statistical inference methods.

Trend-stationary versus difference-stationary processes: why does it matter?* (Optional) I

- Suppose that X_t is a trend-stationary process. For example, suppose it follows the following model:

$$\boxed{X_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)} \tag{13}$$

- To transform X_t into a stationary series we need to remove the linear trend with a regression of X_t on a constant and t.
- The detrended series are the residuals from that regression.
- Suppose now that X_t is a difference-stationary process. For example, suppose it follows a random walk with drift:

$$\boxed{X_t = \beta + X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)} \tag{14}$$

- Here X_t is transformed into a stationary series by first-differencing, ΔX_t.
- The use of the wrong transformation can generate serious problems and lead to misleading inferences for the subsequent econometric analysis.

Trend-stationary versus difference-stationary processes: why does it matter?* (Optional) II

- Suppose that we wrongly apply first differences to a series generated by equation (13). This transformation will introduce artificial autocorrelation in the error term.
- In particular, the transformed series ΔX_t will display a non invertible MA component with a unit root on the MA polynomial.

$$\boxed{\Delta X_t = \beta_1 + \varepsilon_t - \varepsilon_{t-1}}$$

- Suppose now that we wrongly detrend X_t to a series generated by the process (14). In that case the resulting detrended series will be nonstationary or, at least, highly persistent since:

$$\boxed{X_t - E(X_t) = \sum_{i=1}^{t} \varepsilon_i}$$

- Naturally, this is a concern because we don't know in practice if the time series is trend-stationary or difference-stationary and consequently which transformation to use.

Dickey-Fuller test I

- Suppose that X_t follows the process:

$$\boxed{X_t = \phi_1 X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)} \tag{15}$$

- X_t has a unit root if $\phi_1 = 1$ or $\phi(1) = 1 - \phi_1 = 0$. Then the null hypothesis is $H_0 : \phi_1 = 1$ against the stationary alternative $H_1 : -1 < \phi_1 < 1$.
- In practice, it is not realistic to have $\phi_1 \leq -1$ and we will not be worried with this region of ϕ_1. The alternative hypothesis is then stated as $H_1 : \phi_1 < 1$.
- A more convenient formulation is to rewrite the null hypothesis as a function of $\phi(1)$. Subtract both sides of equation (15) by X_{t-1}:

$$\boxed{\Delta X_t = \pi X_{t-1} + \varepsilon_t} \tag{16}$$

where $\pi = \phi_1 - 1 = -\phi(1)$ is the AR polynomial evaluated at $z = 1$.

Dickey-Fuller test II

- The attraction of setting up the model in this way is that this equation format generalizes directly to higher order autoregressive processes (see slides 50 and 51).
- The null and alternative hypothesis are now stated as:

$$H_0 : \pi = 0 \quad \textit{versus} \quad H_1 : \pi < 0$$

- Hence, to apply the test we estimate the statistical model (16) and use the Dickey-Fuller (DF) statistic which is simply the t-ratio of H_0:

$$t_{\widehat{\pi}} = \frac{\widehat{\pi}}{\widehat{\sigma}_{\widehat{\pi}}} \tag{17}$$

  where $\widehat{\pi}$ is the estimate of π obtained from regression (16).
- Notice that X_t has a unit root under H_0. This implies that the test statistic $t_{\widehat{\pi}}$ does not follow any standard distribution such as t or $N(0, 1)$. It follows the so-called **Dickey-Fuller distribution** whose critical values are presented in Table 1.

Dickey-Fuller test III

- Another point worth noticing is the restriction on the deterministic component imposed by the statistical model in (15).
- In particular, X_t is a random walk under the unit root null hypothesis, $H_0 : \pi = 0$:

$$\boxed{X_t = X_{t-1} + \varepsilon_t}$$

  which implies that $E(X_t) = X_0$.
- But under the stationary alternative hypothesis, $H_1 : \pi < 0$, X_t follows an AR(1) process without constant:

$$\boxed{X_t = \phi_1 X_{t-1} + \varepsilon_t, \quad \phi < 1}$$

  which means that $E(X_t) = 0$.
- We conclude that the eventual rejection of H_0 favours a statistical model which restricts the unconditional mean to be equal to 0, $E(X_t) = 0$.
- This fact is rarely observed in a real economic time series data and so this version of the DF statistic is rarely applied in practice.

Dickey-Fuller test with constant I

- In practice, we always include deterministic variables (constant and/or trends) in the statistical model since, in general, $E(X_t) \neq 0$.
- For example, if we add a constant to the process in (16), the Dickey-Fuller regression becomes:

$$\boxed{\Delta X_t = \beta_0 + \pi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)} \tag{18}$$

- As before the unit root null hypothesis and the stationary alternative are defined as $H_0 : \pi = 0$ and $H_1 : \pi < 0$, respectively. The Dickey-Fuller (DF) statistic is obtained as:

$$\boxed{t_{\widehat{\pi}} = \frac{\widehat{\pi}}{\widehat{\sigma}_{\widehat{\pi}}}} \tag{19}$$

where $\widehat{\pi}$ is the estimate of π obtained from regression (18).

Dickey-Fuller test with constant II

- The presence of the constant term in regression (18) changes the asymptotic distribution. The critical values are reported in Table 1.
- As regards to the deterministic components, under the null hypothesis, $H_0 : \pi = 0$, X_t follows a random walk with drift:

$$\boxed{X_t = \beta_0 + X_{t-1} + \varepsilon_t}$$

  which means that we are allowing for the presence of a linear deterministic trend $E(X_t) = X_0 + \beta_0 t$.
- On the other hand, under the stationary alternative, $H_1 : \pi < 0$, X_t follows an AR(1) process with constant:

$$\boxed{X_t = \beta_0 + \phi_1 X_{t-1} + \varepsilon_t, \quad \phi < 1}$$

  which allows the unconditional mean to be different from 0 but restricts to be constant over time $E(X_t) = \frac{\beta_0}{1-\phi_1}$.

Dickey-Fuller test with constant III

- The choice of the appropriate version of the unit root test is very important because the application of the "wrong" test may lead to misleading statistical inference.
- In this case, if the time series is trend-stationary, the $t_{\hat{\pi}}$ statistic is very likely not to reject H_0 and misleads the practitioner to conclude for the presence of a unit root.

Dickey-Fuller test with a linear trend I

- Many times by visual inspection or by economic intuition we may suspect that the time series has a linear time trend.
- In that situation we are doubtful whether X_t has a unit root or is trend-stationary.
- To avoid misleading inferences, we should include a deterministic trend to the Dickey-Fuller regression:

$$\boxed{\Delta X_t = \beta_0 + \beta_1 t + \pi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)} \qquad (20)$$

- The unit root null and the no unit root alternative hypothesis are unchanged and given by $H_0 : \pi = 0$ and $H_1 : \pi < 0$, respectively. The Dickey-Fuller (DF) statistic is obtained as the usual t-ratio:

$$\boxed{t_{\widehat{\pi}} = \frac{\widehat{\pi}}{\widehat{\sigma}_{\widehat{\pi}}}} \qquad (21)$$

where $\widehat{\pi}$ is the estimate of π obtained from regression (20).

Dickey-Fuller test with a linear trend II

- Again the presence of a linear time trend shifts the asymptotic distribution. The critical values are presented in Table 1.
- The presence of a linear trend in the regression (20) again changes the assumption about the unconditional mean of X_t.
- Under the null hypothesis $H_0 : \pi = 0$, X_t follows a random walk with a linear trend. Here the term $\beta_1 t$ is accumulated to produce a quadratic trend in the unconditional mean of X_t:

$$\boxed{X_t = \beta_0 + \beta_1 t + X_{t-1} + \varepsilon_t}$$

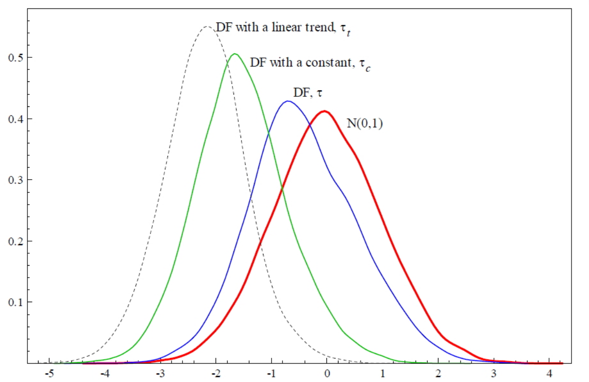
that yields $E(X_t) = \mu_0 + \mu_1 t + \mu_2 t^2$ where μ_0, μ_1 and μ_2 are functions of β_0 and β_1.

- Under the alternative hypothesis, $H_1 : \pi < 0$, X_t obeys an AR(1) process with a linear trend:

$$\boxed{X_t = \beta_0 + \beta_1 t + \phi_1 X_{t-1} + \varepsilon_t, \quad \phi < 1}$$

which delivers a time trend in the mean $E(X_t) = \alpha_0 + \alpha_1 t$ where α_0, α_1 are functions of β_0, β_1 and ϕ.

Dickey-Fuller distributions



Deterministic components in the regression	Significance level		
	10%	5%	1%
No deterministic component	-1.62	-1.94	-2.57
Constant	-2.57	-2.87	-3.44
Constant+trend	-3.13	-3.42	-3.98

Table 1: Critical values from the Dickey-Fuller tests

DF test with no deterministics, intercept or trend and intercept? I

- As argued before the DF statistic with **no deterministics** should not be used in applied work unless there is a very good reason to suspect that the mean of the process X_t is equal to zero, $E(X_t) = 0$.
- The question now is how to decide between **intercept** and **trend+intercept**.
- **Economic reasoning** does not allow some time series to grow continuously over time: unemployment rate, inflation, financial returns, exchange rates, interest rates,. . .. This should be confirmed by **visual inspection** of the graph.

DF test with no deterministics, intercept or trend and intercept? II

- In this case, it **does not** make sense to allow for a **linear time trend** and the DF test with **intercept only** should be applied.
- The same **economic reasoning** and/or **visual inspection** permits that the time series of interest grows over time due to, *e. g.*, inflation or technological progress: variables in nominal terms, prices, consumption, GDP,. . .
- In that case, we should allow for the presence of a **linear trend** and apply the test with **trend and intercept**.

Unit root tests for AR(p) processes I

- All the Dickey -Fuller tests assume that the errors are white-noise $\boxed{\varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)}$.
- Clearly, for many time series 1 lag as in regressions (16), (18) or (20) is insufficient and tipically we include a sufficient number of lags to ensure that $\varepsilon_t \overset{w.n.}{\sim} (0, \sigma_\varepsilon^2)$. In that case, how can we apply the unit root tests?
- For ilustration purposes, consider an AR(2) process without deterministic terms:

$$\boxed{X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t} \tag{22}$$

- The process X_t has a unit root if $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$ has a root at $z = 1$, *i. e.*, if $\phi(1) = 1 - \phi_1 - \phi_2 = 0 \Leftrightarrow \phi_1 + \phi_2 = 1$.
- A possible route would be to estimate the model (22) and test the restriction $\phi_1 + \phi_2 = 1$.

Unit root tests for AR(p) processes II

- However, it is more convenient to rewrite equation (22) such that $\pi = \phi_1 + \phi_2 - 1$ is a parameter of the model. Using $\phi_1 = \pi - \phi_2 + 1$ we have:

$$
\begin{aligned}
X_t =& \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \\
\Leftrightarrow X_t =& \left(\pi - \phi_2 + 1\right) X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \\
\Leftrightarrow X_t - X_{t-1} =& \pi X_{t-1} - \phi_2 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t \\
\Leftrightarrow \Delta X_t =& \pi X_{t-1} + \delta_1 \Delta X_{t-1} + \varepsilon_t
\end{aligned}
$$

  where $\pi = -\phi(1) = \phi_1 + \phi_2 - 1$ and $\delta_1 = -\phi_2$.

- It is possible to show that the general AR(p) process is rewritten as:

$$
\boxed{\Delta X_t = \beta_0 + \beta_1 t + \pi X_{t-1} + \delta_1 \Delta X_{t-1} + \ldots + \delta_{p-1} \Delta X_{t-p+1} + \varepsilon_t}
$$

  where $\pi = -\phi(1) = \phi_1 + \ldots + \phi_p - 1$ and $\delta_1, \ldots, \delta_{p-1}$ are parameters that depend on $\phi_1, \ldots, \phi_p$.

Unit root tests for AR(p) processes III

- X_t has a unit root if $\phi_1 + \ldots + \phi_p = 1$. Hence, the null hypothesis is $H_0 : \pi = 0$ and the no unit root alternative is $H_1 : \pi < 0$.
- The test is again the t-ratio for H_0 and is denoted as **Augmented Dickey-Fuller** test (ADF).
- The critical values are the same as for the Dickey-Fuller test that are reported in Table 1.
- Choice of the number of lags, p:
  1. General-to-specific testing: start with a high lag length value k_{max}. Remove the longest lag if statistical insignificant and reestimate the model. Repeat this process until the longest lags is statistical significant.
  2. Information criteria: AIC, BIC ou HQ.
- For both metholodogies, one should **never forget to verify if the model is well specified**. In particular, one should perform residual diagnostic checks (ACF, Q statistics,...) to ensure that the errors are “white-noise".

Limitations of the ADF tests* (Optional) I

- The unit root tests can have very low power, *i.e.*, the probability of rejecting the unit root null, H_0, when H_0 is true (X_t has a unit root) can be quite low.
- This problem can become particularly serious when:
  1. we have a small sample size ($T \leq 100$).
  2. the inverse roots of the AR polynomialare very close to the unit circle (for example, $\phi_1 = 0.95$ in an AR(1) process). This seems to be a very common phenomena with macroeconomic data.
  3. when irrelevant regressors are included specially, linear trends, slope dummies or to many lags (ΔX_{t-j})
- There is an immense amount of research about unit root tests.
- Nelson & Plosser (1983) was the first published work to study if macroeconomic time series are classified as difference-stationary, trend-stationary or stationary processes.
- They conclude that the most important U.S. macroeconomic time series are well characterized by unit root processes (with or without drift) and so without any mean reverting behaviour.

Limitations of the ADF tests* (Optional) II

Table 5

Tests for autoregressive unit roots[a]

$z_t = \hat{\mu} + \hat{\gamma}t + \hat{\rho}_1 z_{t-1} + \hat{\rho}_2(z_{t-1} - z_{t-2}) + \cdots + \hat{\rho}_k(z_{t-k+1} - z_{t-k}) + \hat{u}_t$.

Series	T	k	$\hat{\mu}$	$t(\hat{\mu})$	$\hat{\gamma}$	$t(\hat{\gamma})$	$\hat{\rho}_1$	$\tau(\hat{\rho}_1)$	$s(\hat{u})$	r_1
Real GNP	62	2	0.819	3.03	0.006	3.03	0.825	−2.99	0.058	−0.02
Nominal GNP	62	2	1.06	2.37	0.006	2.34	0.899	−2.32	0.087	0.03
Real per capita GNP	62	2	1.28	3.05	0.004	3.01	0.818	−3.04	0.059	−0.02
Industrial production	111	6	0.103	4.32	0.007	2.44	0.835	−2.53	0.097	0.03
Employment	81	3	1.42	2.68	0.002	2.54	0.861	−2.66	0.035	0.10
Unemployment rate	81	4	0.513	2.81	−0.000	−0.23	0.706	−3.55*	0.407	0.02
GNP deflator	82	2	0.260	2.55	0.002	2.65	0.915	−2.52	0.046	−0.03
Consumer prices	111	4	0.090	1.76	0.001	2.84	0.986	−1.97	0.042	−0.06
Wages	71	3	0.566	2.30	0.004	2.30	0.910	−2.09	0.060	0.00
Real wages	71	2	0.487	3.10	0.004	3.14	0.831	−3.04	0.034	−0.01
Money stock	82	2	0.133	3.52	0.005	3.03	0.916	−3.08	0.047	0.03
Velocity	102	1	0.052	0.99	−0.000	−0.65	0.941	−1.66	0.067	0.11
Interest rate	71	3	−0.186	−0.95	0.003	1.75	1.03	0.686	0.283	−0.02
Common stock prices	100	3	0.481	2.02	0.003	2.37	0.913	−2.05	0.158	0.20

[a] z_t represents the natural logs of annual data except for the bond yield. $t(\hat{\mu})$ and $t(\hat{\gamma})$ are the ratios of the OLS estimates of μ and γ to their respective standard errors. $\tau(\hat{\rho}_1)$ is the ratio of $\hat{\rho}_1 - 1$ to its standard error. $s(\hat{u})$ is the standard error of the regression and r_1 is the first-order autocorrelation coefficient of the residuals. The values of $\tau(\hat{\rho}_1)$ denoted by an (*) are smaller than the 0.05 one tail critical value of the distribution of $\tau(\hat{\rho}_1)$ and similarly for $\hat{\rho}_1$. It should also be noted that $t(\hat{\mu})$ and $t(\hat{\gamma})$ are not distributed as normal random variables.

- According to Google Scholar, this paper has been quoted 4783 vezes until today.
- A high proportion of these citations represent new unit root tests that contradict or reinforce the NP conclusions.

Limitations of the ADF tests* (Optional) III

- This fact shows that the unit root inference problem is very important but also very complicated or even impossible to solve. The most important developments regarding more powerful and robust unit root tests are the following (available in EViews):
  1. ERS Tests developed by Elliot, Rothemberg e Stock (1996).
  2. NG and Perron tests developed by Serena Ng and Pierre Perron (2001).

Bibliography

- Enders, W. (2009). Applied Econometric Time Series, chapters 4.1-4.3, 4.5-4.8
- Tsay, R. S. (2005). Analysis of Financial Time Series, chapter 2.7
- Wei, W. W. S. (2005). Time series analysis: Univariate and Multivariate Methods, chapter 4
- Mills, T. C. (1991). Time series techniques for economists, chapter 6