Econometrics 2 – Fall 2005

Non-Stationary Time Series and Unit Root Tests

Heino Bohn Nielsen

Introduction

- Many economic time series are trending.
- Important to distinguish between two important cases:
  (1) A stationary process with a deterministic trend:
  Shocks have transitory effects.
  (2) A process with a stochastic trend or a unit root:
  Shocks have permanent effects.
- Why are unit roots important?
  (1) Interesting to know if shocks have permanent or transitory effects.
  (2) It is important for forecasting to know if the process has an attractor.
  (3) Stationarity was required to get a LLN and a CLT to hold.
  For unit root processes, many asymptotic distributions change!
  Later we look at regressions involving unit root processes: spurious regression and cointegration.

Outline of the Lecture

(1) Difference between trend stationarity and unit root processes.

(2) Unit root testing.

(3) Dickey-Fuller test.

(4) Caution on deterministic terms.

(5) An alternative test (KPSS).

Trend Stationarity

- Consider a stationary AR(1) model with a deterministic linear trend term:

$$Y_t = \theta Y_{t-1} + \delta + \gamma t + \epsilon_t, \quad t = 1, 2, ..., T, \qquad (*)$$

where $|\theta| < 1$, and Y_0 is an initial value.

- The solution for Y_t (MA-representation) has the form

$$Y_t = \theta^t Y_0 + \mu + \mu_1 t + \epsilon_t + \theta \epsilon_{t-1} + \theta^2 \epsilon_{t-2} + \theta^3 \epsilon_{t-3} + ...$$

Note that the mean,

$$E[Y_t] = \theta^t Y_0 + \mu + \mu_1 t \to \mu + \mu_1 t \quad \text{for} \quad T \to \infty,$$

contains a linear trend, while the variance is constant:

$$V[Y_t] = V[\epsilon_t + \theta \epsilon_{t-1} + \theta^2 \epsilon_{t-2} + ...] = \sigma^2 + \theta^2 \sigma^2 + \theta^4 \sigma^2 + ... = \frac{\sigma^2}{1 - \theta^2}.$$

- The original process, Y_t, is not stationary.
  The deviation from the mean,

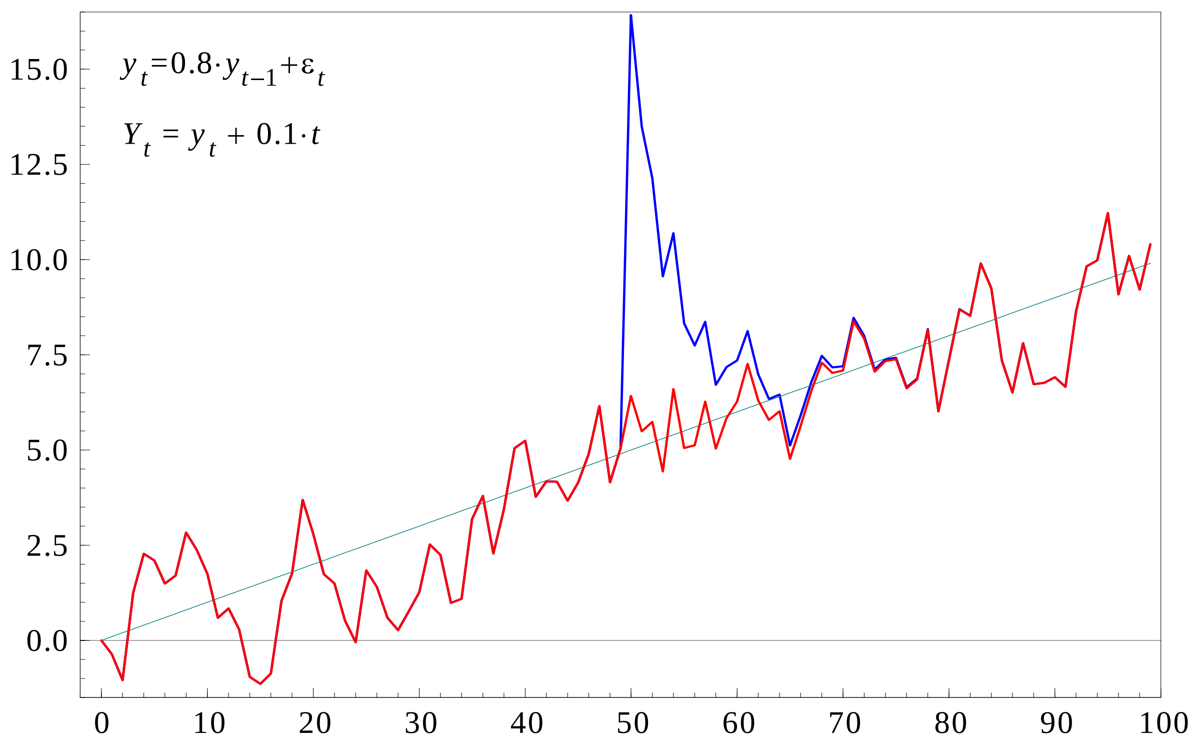
$$y_t = Y_t - E[Y_t] = Y_t - \mu - \mu_1 t$$

  is a stationary process. The process Y_t is called trend-stationary.

- The stochastic part of the process is stationary and shocks have transitory effects
  We say that the process is mean reverting.
  Also, we say that the process has an attractor, namely the mean, $\mu + \mu_1 t$.

- We can analyze deviations from the mean, y_t.
  From the Frisch-Waugh theorem this is the same as a regression including a trend.

Shock to a Trend-Stationarity Process



Unit Root Processes

- Consider the AR(1) model with a unit root, $\theta = 1$:

$$\begin{aligned} Y_t &= Y_{t-1} + \delta + \epsilon_t, \quad t = 1, 2, ..., T, \\ \text{or} \quad \Delta Y_t &= \delta + \epsilon_t, \end{aligned} \qquad (**)$$

  where Y_0 is the initial value.

- Note that $z = 1$ is a root in the autoregressive polynomial, $\theta(L) = (1 - L)$. $\theta(L)$ is not invertible and Y_t is non-stationary.

- The process ΔY_t is stationary. We denote Y_t a difference stationary process.

- If ΔY_t is stationary while Y_t is not, Y_t is called integrated of first order, I(1). A process is integrated of order d, I(d), if it contains d unit roots.

- The solution for Y_t is given by

$$Y_t = Y_0 + \sum_{i=1}^{t} \Delta Y_i = Y_0 + \sum_{i=1}^{t} (\delta + \epsilon_i) = Y_0 + \delta t + \sum_{i=1}^{t} \epsilon_i,$$

  with moments

$$E[Y_t] = Y_0 + \delta t \quad \text{and} \quad V[Y_t] = t \cdot \sigma^2$$

Remarks:

(1) The effect of the initial value, Y_0, stays in the process.

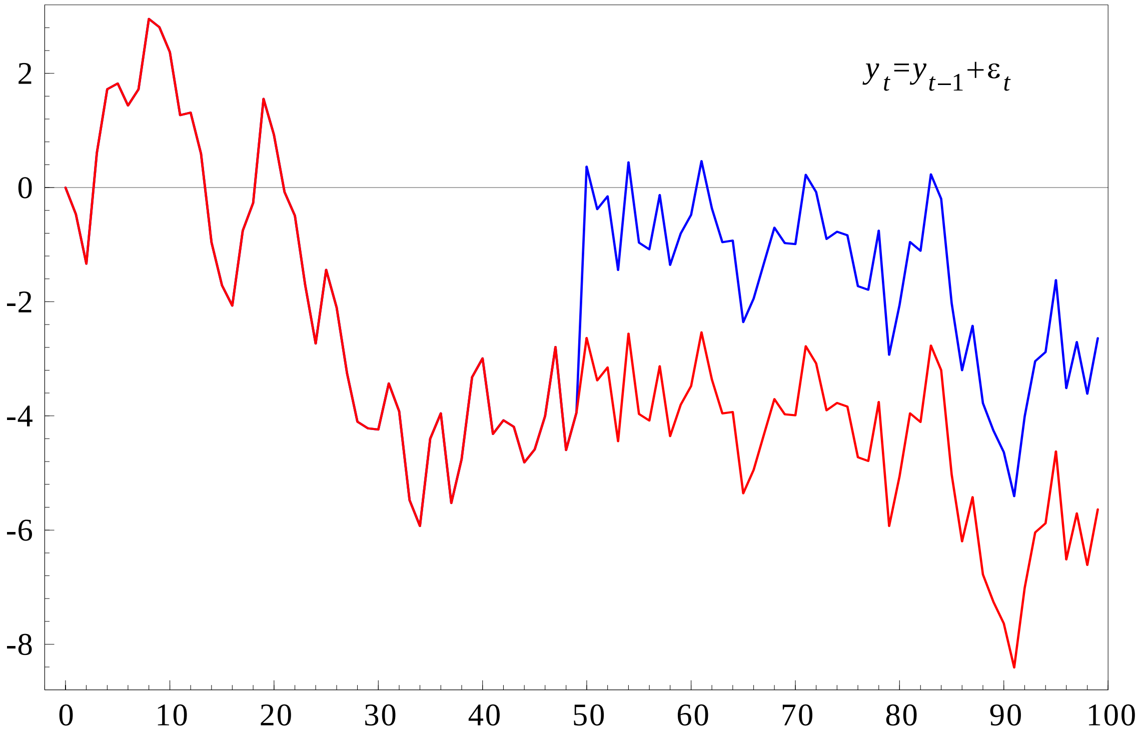
(2) The innovations, ϵ_t, are accumulated to a random walk, $\sum \epsilon_i$.
This is denoted a stochastic trend.
Note that shocks have permanent effects.

(3) The constant δ is accumulated to a linear trend in Y_t.
The process in $(**)$ is denoted a random walk with drift.

(4) The variance of Y_t grows with t.

(5) The process has no attractor.

Shock to a Unit Root Process



Unit Root Tests

- A good way to think about unit root tests:

  We compare two relevant models: H_0 and H_A.

  (1) What are the properties of the two models?

  (2) Do they adequately describe the data?

  (3) Which one is the null hypothesis?

- Consider two alternative test:

  (1) Dickey-Fuller test: H_0 is a unit root, H_A is stationarity.

  (2) KPSS test: H_0 is stationarity, H_A is a unit root.

- Often difficult to distinguish in practice (Unit root tests have low power).

  Many economic time series are persistent, but is the root 0.95 or 1.0?

The Dickey-Fuller (DF) Test

- Idea: Set up an autoregressive model for y_t and test if $\theta(1) = 0$.

- Consider the AR(1) regression model

$$y_t = \theta y_{t-1} + \epsilon_t.$$

The unit root null hypothesis against the stationary alternative corresponds to

$$H_0 : \theta = 1 \quad \text{against} \quad H_A : \theta < 1.$$

- Alternatively, the model can be formulated as

$$\Delta y_t = (\theta - 1) y_{t-1} + \epsilon_t = \pi y_{t-1} + \epsilon_t,$$

where $\pi = \theta - 1 = \theta(1)$. The unit root hypothesis translates into

$$H_0 : \pi = 0 \quad \text{against} \quad H_A : \pi < 0.$$

- The Dickey-Fuller (DF) test is simply the $t-$ test for H_0 :

$$\widehat{\tau} = \frac{\widehat{\theta} - 1}{\text{se}(\widehat{\theta})} = \frac{\widehat{\pi}}{\text{se}(\widehat{\pi})}.$$

The asymptotic distribution of $\widehat{\tau}$ is not normal!
The distribution depends on the deterministic components.
In the simple case, the 5% critical value (one-sided) is -1.95 and not -1.65.

Remarks:

(1) The distribution only holds if the errors ϵ_t are IID (check that!)
If autocorrelation, allow more lags.

(2) In most cases, MA components are approximated by AR lags.
The distribution for the test of $\theta(1) = 0$ also holds in an ARMA model.

Augmented Dickey-Fuller (ADF) test

- The DF test is extended to an AR(p) model. Consider an AR(3):

$$y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + \theta_3 y_{t-3} + \epsilon_t.$$

A unit root in $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3$ corresponds to $\theta(1) = 0$.

- The test is most easily performed by rewriting the model:

$$\begin{aligned}
y_t - y_{t-1} &= (\theta_1 - 1)y_{t-1} + \theta_2 y_{t-2} + \theta_3 y_{t-3} + \epsilon_t \\
y_t - y_{t-1} &= (\theta_1 - 1)y_{t-1} + (\theta_2 + \theta_3)y_{t-2} + \theta_3(y_{t-3} - y_{t-2}) + \epsilon_t \\
y_t - y_{t-1} &= (\theta_1 + \theta_2 + \theta_3 - 1)y_{t-1} + (\theta_2 + \theta_3)(y_{t-2} - y_{t-1}) + \theta_3(y_{t-3} - y_{t-2}) + \epsilon_t \\
\Delta y_t &= \pi y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + \epsilon_t,
\end{aligned}$$

where

$$\begin{aligned}
\pi &= \theta_1 + \theta_2 + \theta_3 - 1 = -\theta(1) \\
c_1 &= -(\theta_2 + \theta_3) \\
c_2 &= -\theta_3.
\end{aligned}$$

- The hypothesis $\theta(1) = 0$ again corresponds to

$$H_0 : \pi = 0 \quad \text{against} \quad H_A : \pi < 0.$$

The $t-$test for H_0 is denoted the augmented Dickey-Fuller (ADF) test.

- To determine the number of lags, k, we can use the normal procedures.
  – General-to-specific testing: Start with $k_{\max}$ and delete insignificant lags.
  – Estimate possible models and use information criteria.
  – Make sure there is no autocorrelation.

- Verbeek suggests to calculate the DF test for all values of k.
  This is a robustness check, but be careful!
  Why would we look at tests based on inferior/misspecified models?

- An alternative to the ADF test is to correct the DF test for autocorrelation.
  Phillips-Perron non-parametric correction based on HAC standard errors.
  Quite complicated and likely to be inferior in small samples.

Deterministic Terms in the DF Test

- The deterministic specification is important:
  - We want an adequate model for the data.
  - The deterministic specification changes the asymptotic distribution.

- If the variable has a non-zero level, consider a regression model of the form

$$\Delta Y_t = \pi Y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + \delta + \epsilon_t.$$

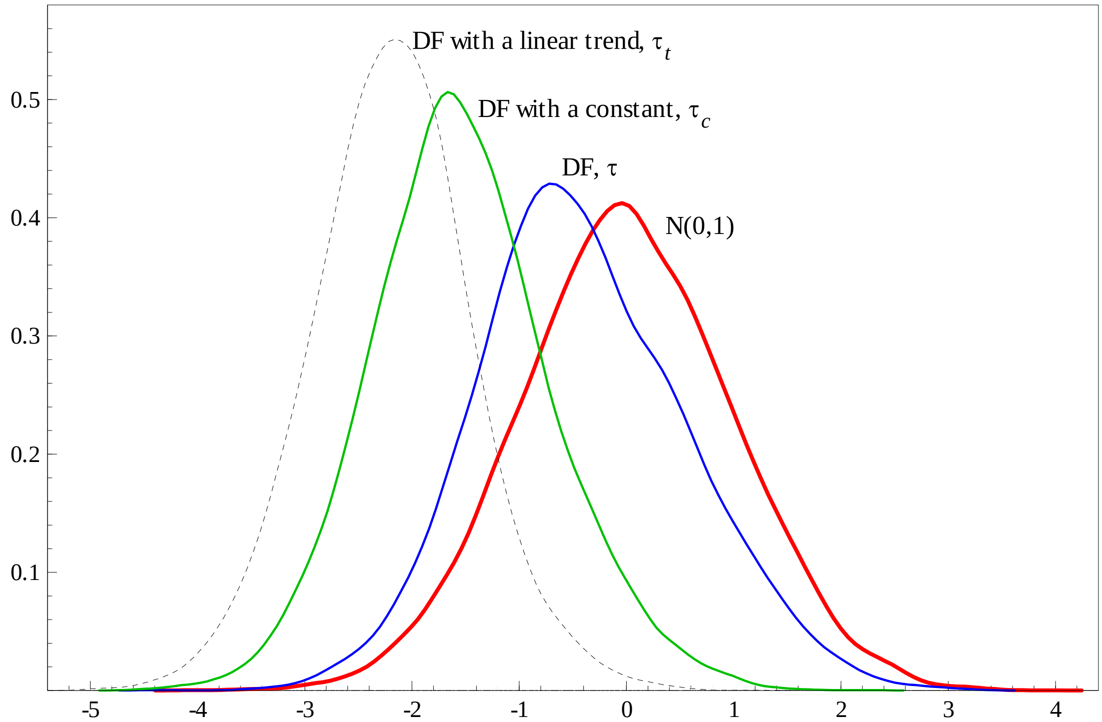
The ADF test is the $t-$test, $\widehat{\tau}_c = \widehat{\pi}/\text{se}(\widehat{\pi})$.
The critical value at a 5% level is -2.86.

- If the variable has a deterministic trend, consider a regression model of the form

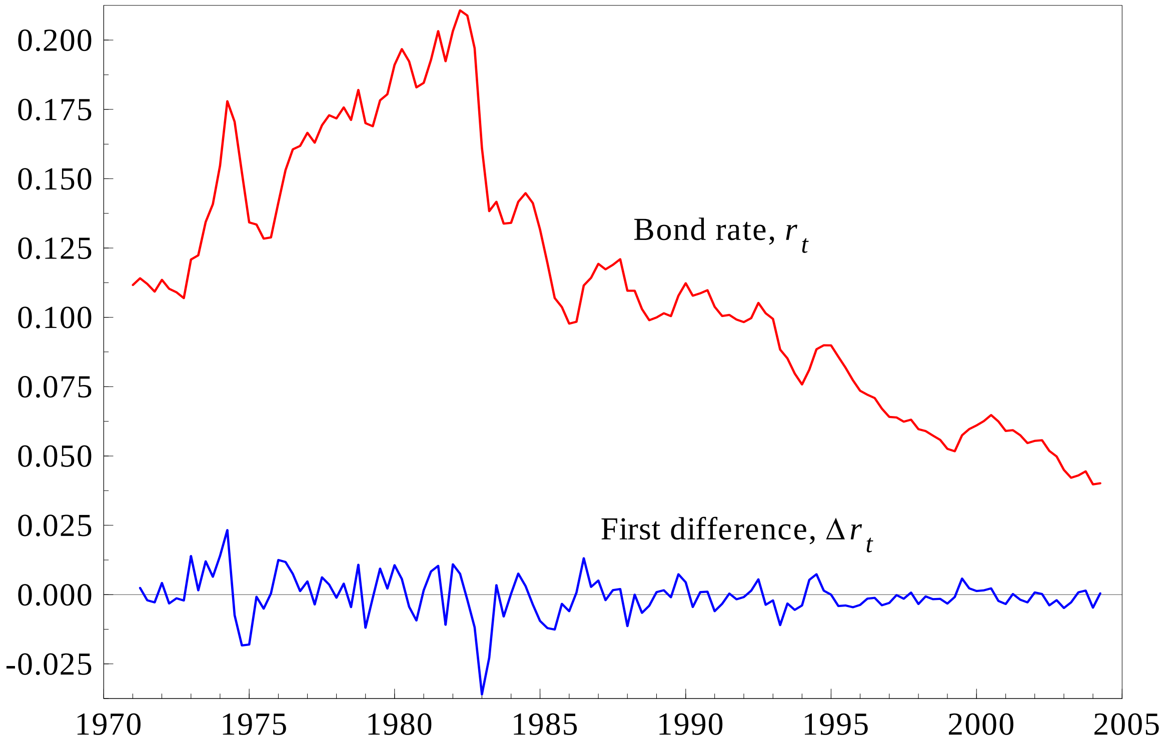
$$\Delta Y_t = \pi Y_{t-1} + c_1 \Delta y_{t-1} + c_2 \Delta y_{t-2} + \delta + \gamma t + \epsilon_t.$$

The ADF test is the $t-$test, $\widehat{\tau}_t = \widehat{\pi}/\text{se}(\widehat{\pi})$.
The critical value at a 5% level is -3.41.

The DF Distributions



Empirical Example: Danish Bond Rate



- An AR(4) model gives

$$\Delta r_t = \underset{(-0.672)}{-0.0093 r_{t-1}} + \underset{(4.49)}{0.4033 \Delta r_{t-1}} \underset{(-0.199)}{-0.0192} \Delta r_{t-2} \underset{(-0.817)}{-0.0741} \Delta r_{t-3} + \underset{(0.406)}{0.0007}.$$

  Removing insignificant terms produce a model

$$\Delta r_t = \underset{(-0.911)}{-0.0122 r_{t-1}} + \underset{(4.70)}{0.3916 \Delta r_{t-1}} + \underset{(0.647)}{0.0011}.$$

  The 5% critical value ($T = 100$) is -2.89, so we do not reject the null of a unit root.

- We can also test for a unit root in the first difference.
  Deleting insignificant terms we find a preferred model

$$\Delta^2 r_t = \underset{(-7.49)}{-0.6193 \Delta r_{t-1}} \underset{(-0.534)}{-0.00033}.$$

  Here we safely reject the null hypothesis of a unit root $(-7.49 \ll -2.89)$.

- Based on the test we conclude that r_t is non-stationary while Δr_t is stationary.
  That is $r_t \sim$I(1)

A Note of Caution on Deterministic Terms

- The way to think about the inclusion of deterministic terms is via the factor representation:

$$y_t = \theta y_{t-1} + \epsilon_t$$
$$Y_t = y_t + \mu$$

It follows that

$$\begin{aligned} (Y_t - \mu) &= \theta(Y_{t-1} - \mu) + \epsilon_t \\ Y_t &= \theta Y_{t-1} + (1-\theta)\mu + \epsilon_t \\ Y_t &= \theta Y_{t-1} + \delta + \epsilon_t \end{aligned}$$

which implies a common factor restriction.

- If $\theta = 1$, then implicitly the constant should also be zero, i.e.

$$\delta = (1-\theta)\mu = 0.$$

- The common factor is not imposed by the normal $t-$test. Consider

$$Y_t = \theta Y_{t-1} + \delta + \epsilon_t.$$

The hypotheses

$$H_0 : \theta = 1 \quad \text{against} \quad H_A : \theta < 1,$$

imply

$$\begin{aligned} H_A &: Y_t = \mu + \text{stationary process} \\ H_0 &: Y_t = Y_0 + \delta t + \text{stochastic trend}. \end{aligned}$$

- We compare a model with a linear trend against a model with a non-zero level! Potentially difficult to interpret.

- A simple alternative is to consider the combined hypothesis

$$H_0^* : \pi = \delta = 0.$$

- The hypothesis H_0^* can be tested by running the two regressions

$$\begin{aligned} H_A &: \Delta Y_t = \pi Y_{t-1} + \delta + \epsilon_t \\ H_0^* &: \Delta Y_t = \epsilon_t, \end{aligned}$$

and perform a likelihood ratio test

$$\tau_{LR} = T \cdot \log \left(\frac{\text{RSS}_0}{\text{RSS}_A} \right) = -2 \left(\text{loglik}_0 - \text{loglik}_A \right),$$

where RSS_0 and RSS_A denote the residual sum of squares.
The 5% critical value is 9.13.

Same Point with a Trend

- The same point could be made with a trend term

$$\Delta Y_t = \pi Y_{t-1} + \delta + \gamma t + \epsilon_t.$$

Here, the common factor restriction implies that if $\pi = 0$ then $\gamma = 0$.

- Since we do not impose the restriction under the null, the trend will accumulate.
A quadratic trend is allowed under H_0, but only a linear trend under H_A.

- A solution is to impose the combined hypothesis

$$H_0^* : \pi = \gamma = 0.$$

This is done by running the two regressions

$$\begin{aligned} H_A &: \Delta Y_t = \pi Y_{t-1} + \delta + \gamma t + \epsilon_t \\ H_0^* &: \Delta Y_t = \delta + \epsilon_t, \end{aligned}$$

and perform a likelihood ratio test. The 5% critical value for this test is 12.39.

Special Events

- Unit root tests assess whether shocks have transitory or permanent effects.
  The conclusions are sensitive to a few large shocks.

- Consider a one-time change in the mean of the series, a so-called break.
  This is one large shock with a permanent effect.
  Even if the series is stationary, such that normal shocks have transitory effects, the presence of a break will make it look like the shocks have permanent effects.
  That may bias the conclusion towards a unit root.

- Consider a few large outliers, i.e. a single strange observations.
  The series may look more mean reverting than it actually is.
  That may bias the results towards stationarity.

A Reversed Test: KPSS

- Sometimes it is convenient to have stationarity as the null hypothesis.
  KPSS (Kwiatkowski, Phillips, Schmidt, and Shin) Test.

- Assume there is no trend. The point of departure is a DGP of the form

$$Y_t = \xi_t + e_t,$$

  where e_t is stationary and ξ_t is a random walk, i.e.

$$\xi_t = \xi_{t-1} + v_t, \quad v_t \sim IID(0, \sigma_v^2).$$

  If the variance is zero, $\sigma_v^2 = 0$, then $\xi_t = \xi_0$ for all t and Y_t is stationary.
  Use a simple regression:

$$Y_t = \widehat{\mu} + \widehat{e}_t, \qquad (*)$$

  to find the estimated stochastic component. Under the null, $\widehat{e}_t$ is stationary.

- That observation can be used to design a test:

$$H_0 : \sigma_v^2 = 0 \quad \text{against} \quad H_A : \sigma_v^2 > 0.$$

- The test statistic is given by

$$KPSS = \frac{1}{T^2} \cdot \frac{\sum_{t=1}^{T} S_t^2}{\widehat{\sigma}_\infty^2},$$

where $S_t = \sum_{s=1}^{t} \widehat{e}_s$ is a partial sum; $\widehat{\sigma}_\infty^2$ is a HAC estimator of the variance of $\widehat{e}_t$. (This is an LM test for constant parameters against a RW parameter).

- The regression in $(*)$ can be augmented with a linear trend. Critical values:

Deterministic terms in regression $(*)$	Critical values		
	0.10	0.05	0.01
Constant	0.347	0.463	0.739
Constant and trend	0.119	0.146	0.216

- Can be used for confirmatory analysis:

		KPSS: Rejection of I(0)	KPSS: Non-rejection of I(0)
DF:	Rejection of I(1)	?	I(0)
	Non-rejection of I(1)	I(1)	?