Eco 5316 Time Series Econometrics

Lecture 7 Nonstationary Time Series

Nonstationary Time Series

a lot of time series in economics and finance are not weakly stationary and instead

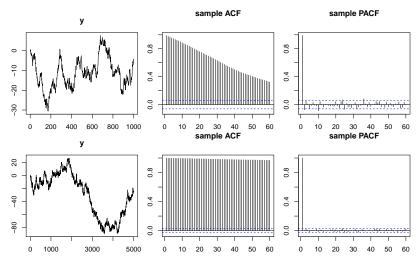
- ▶ show linear or exponential trend
- show stochastic trend grow or fall over time or meander without a constant long-run mean
- show increasing variance over time

examples

- ► GDP, consumption, investment, exports, imports, ...
- ▶ industrial production, retail sales, ...
- interest rates, foreign exchange rates, stock market indices, prices of commodities....
- ▶ unemployment rate, labor force participation rate, . . .
- ► loans, federal debt, ...

Nonstationary Time Series

A very slowly decaying ACF suggests nonstationarity and presence of deterministic or stochastic trend in the time series, e.g. for $y_t=y_{t-1}+\varepsilon_t$



Transformations

Detrending - regressing y_t on intercept and time trend - proper treatment id $\{y_t\}$ is trend stationary

Differencing - proper treatment if $\{y_t\}$ is difference stationary

Log transformation and differencing - proper treatment if $\{y_t\}$ grows exponentially and shows increasing variability over time

Trend-Stationary Time Series

lacktriangleright consider times series $\{y_t\}$ that follows

$$y_t = \alpha + \mu t + \varepsilon_t$$

where ε_t is a weakly stationary time series

- $E(y_t) = \alpha + \mu t$ and $var(y_t) = var(\varepsilon_t) = const.$
- ▶ since $E(y_t) \neq const.$ time series $\{y_t\}$ is not weakly stationary
- ▶ $\{y_t\}$ can however be made stationary by removing time trend using a regression of y_t on constant and time
- $\{y_t\}$ is **trend stationary** time series

Difference-Stationary Time Series

Random Walk

• suppose ε_t is white noise, consider a version of AR(1) model with $\phi_0=0$ and $\phi_1=1$

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

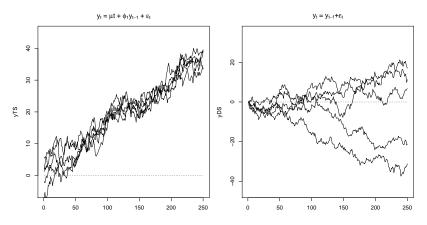
or, by repeated substitution

$$y_t = \alpha + \sum_{j=1}^t \varepsilon_j$$

where $\alpha = y_0$

- $E(y_t) = \alpha$ and $var(y_t) = var(\sum_{j=1}^t \varepsilon_j) = t\sigma_{\varepsilon}^2$
- ▶ since $var(y_t) \neq const.$ time series $\{y_t\}$ is not weakly stationary
- $\{y_t\}$ can not be made difference stationary by removing time trend using a regression of y_t on constant and time
- $lackbox \{y_t\}$ can however be made stationary by differencing
- $ightharpoonup \{y_t\}$ is difference stationary time series

five simulations of trend stationary time series vs random walk



Difference-Stationary Time Series

Random Walk with Drift

▶ suppose ε_t is white noise, consider a version of AR(1) model with $\phi_1 = 1$

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

and by repeated substitution

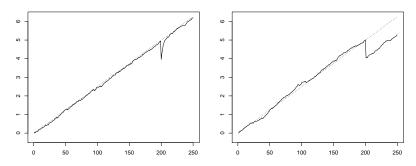
$$y_t = \alpha + \mu t + \sum_{j=1}^t \varepsilon_j$$

where $\alpha = y_0$

- ► $E(y_t) = \alpha + \mu t$ and $var(y_t) = var(\sum_{j=1}^t \varepsilon_j) = t\sigma_{\varepsilon}^2$
- ▶ $E(y_t) \neq const.$ and $var(y_t) \neq const.$ so $\{y_t\}$ is not weakly stationary
- $\{y_t\}$ can not be made difference stationary by removing time trend using a regression of y_t on constant and time
- $lackbox \{y_t\}$ can however be made stationary by differencing
- ▶ $\{y_t\}$ is **difference stationary** time series

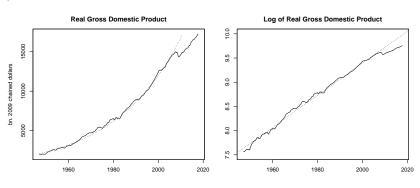
It is important to be able to distinguish between the two cases:

- ▶ with trend stationary series shocks have transitory effects
- ▶ with difference stationary series shocks have **permanent effects**



In addition, as we will see later additional issues arise with difference stationary series in the context of multivariate time series analysis

U.S. GDP and the effect of 2008-2009 recession permanent effect or structural break?



Unit-root Time Series

Autoregressive Integrated Moving-Average (ARIMA) Models

- non-stationary time series is said to contain a unit root or to be integrated of order one, I(1), if it can be made stationary by applying first differences
- ▶ time series $\{y_t\}$ follows an ARIMA(p,1,q) process if $\Delta y_t = (1-L)y_t$ follows a stationary and invertible ARMA(p,q) process, so that

$$\phi(L)(1-L)y_t = \mu + \theta(L)\varepsilon_t$$

Unit-root Time Series

Autoregressive Integrated Moving-Average (ARIMA) Models

- ▶ non-stationary time series is said to be **integrated of order** d, I(d), if it can be made stationary by differencing d times
- ▶ time series $\{y_t\}$ follows an ARIMA(p,d,q) process if $\Delta^d y_t = (1-L)^d y_t$ follows a stationary and invertible ARMA(p,q) process, thus

$$\phi(L)(1-L)^d y_t = \mu + \theta(L)\varepsilon_t$$

 \blacktriangleright note that pure random walk and random walk with drift are special cases, an $\mathsf{ARIMA}(0,1,0)$

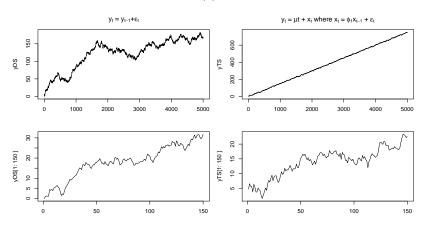
$$(1-L)y_t = \mu + \varepsilon_t$$

with $\mu=0$ in case of pure random walk and $\mu\neq 0$ in case of random walk with drift

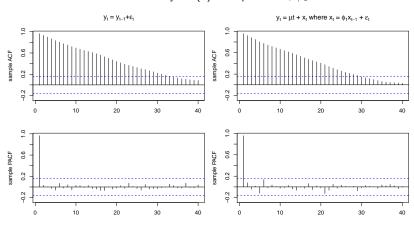
it is often very hard to distinguish random walk and trend stationary model:

150 vs 5000 observations of

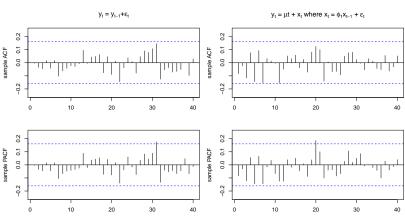
random walk vs. trend stationary AR(1) with $\mu =$ 0.15, $\phi_1 =$ 0.95



ACF and PACF for 150 observations of y_t under random walk vs. trend stationary AR(1) with $\mu=$ 0.15, $\phi_1=$ 0.95



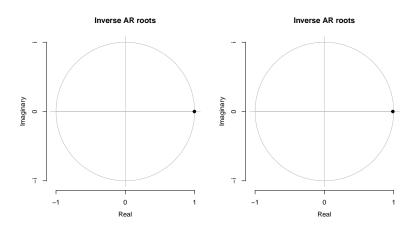
ACF and PACF for 150 observations of first difference Δy_t under random walk vs. trend stationary AR(1) with $\mu=0.15,\,\phi_1=0.95$



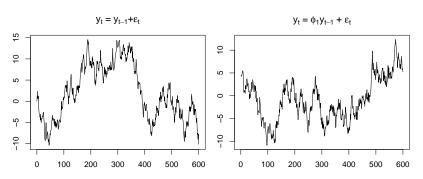
random walk vs. trend stationary AR(1) with $\mu=$ 0.15, $\phi_1=$ 0.95

```
## Series: yDS[1:T]
## ARIMA(1.0.0) with non-zero mean
##
## Coefficients:
##
           ar1
                  mean
##
        0.9971 16.279
## s.e. 0.0038 12.711
##
## sigma^2 estimated as 1.138: log likelihood=-224.1
## ATC=454.19 ATCc=454.36 BTC=463.22
## Series: vTS[1:T]
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##
           ar1
                   mean
        0.9878 13.7733
## s.e. 0.0123 4.7683
##
## sigma^2 estimated as 1.065: log likelihood=-218.44
## AIC=442.87 AICc=443.04 BIC=451.91
```

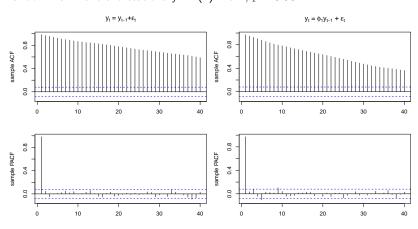
random walk vs. trend stationary AR(1) with $\mu=$ 0.15, $\phi_1=$ 0.95



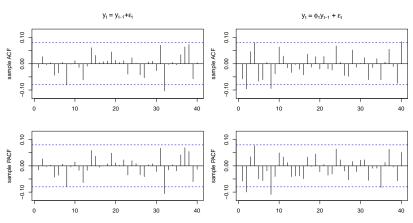
also very hard to distinguish random walk model and highly persistent AR(1): random walk I(1) vs. AR(1) with $\phi_1=0.98$



ACF and PACF for y_t under random walk vs. trend stationary AR(1) with $\phi_1=0.98$



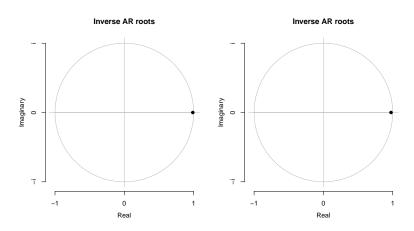
ACF and PACF for first difference Δy_t under random walk vs. trend stationary AR(1) with $\phi_1=$ 0.98



random walk vs. trend stationary AR(1) with $\phi_1=0.98$

```
## Series: vI1
## ARIMA(1.0.0) with non-zero mean
##
## Coefficients:
##
           ar1
                  mean
##
        0.9885 0.4748
## s.e. 0.0060 3.2424
##
## sigma^2 estimated as 1.034: log likelihood=-863.67
## ATC=1733.33 ATCc=1733.37 BTC=1746.53
## Series: vAR1
## ARIMA(1,0,0) with non-zero mean
##
## Coefficients:
##
           ar1
                   mean
        0.9760 -0.2034
##
## s.e. 0.0087 1.6538
##
## sigma^2 estimated as 1.054: log likelihood=-867.77
## AIC=1741.55 AICc=1741.59 BIC=1754.74
```

random walk vs. trend stationary AR(1) with $\mu =$ 0.15, $\phi_1 =$ 0.98



- two types of tests for nonstationarity
 - lacktriangle unit root tests: H_0 is difference stationarity, H_A is trend stationarity
 - ightharpoonup stationarity tests: H_0 is trend stationary, H_A is difference stationarity
- lacktriangle in general, the approach of these tests is to consider $\{y_t\}$ as a sum

$$y_t = d_t + z_t + \varepsilon_t$$

where d_t is a deterministic component (time trend, seasonal component, etc.), z_t is a stochastic trend component and ε_t is a stationary process

 \blacktriangleright tests then investigate whether z_t is present

Augmented Dickey-Fuller (ADF) test

▶ main idea: suppose $\{y_t\}$ follows AR(1)

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

then

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

where
$$\gamma = \phi_1 - 1$$

• if $\{y_t\}$ is I(1) then $\gamma=0$, otherwise $\gamma<0$

Augmented Dickey-Fuller (ADF) test

• unit root test H_0 : time series $\{y_t\}$ has a unit root H_A : time series $\{y_t\}$ is stationary (with zero mean - model A), level stationary (with non-zero mean - model B) or trend stationary (stationary around a deterministic trend - model C)

$$\begin{array}{ll} \operatorname{model} \; \mathsf{A} & \Delta y_t = \gamma y_{t-1} + \sum_{i=1}^{p-1} \rho_i \Delta y_{t-i} + e_t \\ \\ \operatorname{model} \; \mathsf{B} & \Delta y_t = \gamma y_{t-1} + \mu + \sum_{i=1}^{p-1} \rho_i \Delta y_{t-i} + e_t \\ \\ \operatorname{model} \; \mathsf{C} & \Delta y_t = \gamma y_{t-1} + \mu + \beta t + \sum_{i=1}^{p-1} \rho_i \Delta y_{t-i} + e_t \end{array}$$

- lacktriangleright if $\{y_t\}$ contains a unit root/is difference stationary, $\hat{\gamma}$ will be insignificant
- ▶ test $H_0: \gamma = 0$ against $H_A: \gamma < 0$; if t-statistics for γ is lower than critical values we reject the null hypothesis of a unit root (one-sided left-tailed test)

Augmented Dickey-Fuller (ADF) test

If $\gamma < 0$ then

- \blacktriangleright under model A y_t fluctuates around zero
- lacktriangle under model B if $\mu \neq 0$ then y_t fluctuates around a non-zero mean
- under model C if $\mu \neq 0$, $\beta \neq 0$ then y_t fluctuates around linear deterministic trend βt

If $\gamma=0$ then

- lacktriangle under model A y_t contains stochastic trend only
- \blacktriangleright under model B if $\mu \neq 0$ then y_t contains both a linear deterministic trend μt and a stochastic trend
- under model C if $\mu \neq 0$, $\beta \neq 0$ then y_t contains a quadratic deterministic trend βt^2 and a stochastic trend

Augmented Dickey-Fuller (ADF) test

- ▶ lags Δy_{t-i} used in the test are in order to control for the possible higher order autocorrelation
- ▶ number of lags can be chosen by a simple procedure: start with some reasonably large number of lags p_{max} and check the significance of the coefficient on the highest lag with a t-test; if insignificant at the 10 % level, reduce the number of lags by one, proceed in this way until achieving significance
- lacktriangle an alternative approach: select the number of lags p to minimize AIC or BIC
- ightharpoonup if p is too small errors will be serially correlated which will bias the test, if p is too large power of the test will suffer
- ▶ it is better to err on the side of including too many lags
- ightharpoonup ADF has very low power against I(0) alternatives that are close to being I(1), it can't distinguish highly persistent stationary processes from nonstationary processes well

Augmented Dickey-Fuller (ADF) test

- including constant and trend in the regression also weakens the test (model C is thus the weakest on, model A the strongest one)
- ▶ if possible, we want to exclude the constant and/or the trend, but if they are incorrectly excluded, the test will be biased
- in addition to providing critical values to testing whether $\gamma=0$, Dickey and Fuller also provide critical values for the following three F tests:
 - ϕ_1 statistic for model B to test $H_0: \gamma = \mu = 0$
 - ϕ_2 statistic for model C to test $H_0: \gamma = \mu = \beta = 0$
 - ϕ_3 statistic for model C to test $H_0: \gamma = \beta = 0$
- these allow us to test whether we can restrict the test

Proposed Full Procedure for ADF test

- **Step 1.** estimate model C and use τ_3 statistic to test H_0 : $\gamma = 0$
- if H_0 can not be rejected continue to Step 2
- if H_0 is rejected conclude that y_t is trend stationary

Step 2. use ϕ_3 statistic to test H_0 : $\gamma = \beta = 0$

- if H_0 can not be rejected continue to step 3
- if H_0 is rejected estimate restricted model $\Delta y_t = \mu + \beta t + \sum_{i=1}^{p-1} \rho_i \Delta y_{t-i} + e_t$ and use t statistic to test $H_0: \beta = 0$
 - if H_0 can not be rejected continue to Step 3
 - if H_0 is rejected conclude that y_t is difference stationary with quadratic trend

Step 3. estimate model B and use τ_2 statistic to test H_0 : $\gamma=0$

- if H_0 can not be rejected continue to Step 4
- if H_0 is rejected conclude that y_t is trend stationary

Step 4. use ϕ_1 statistic to test H_0 : $\gamma = \mu = 0$

- if H_0 can not be rejected continue to step 5
- if H_0 is rejected estimate restricted model $\Delta y_t = \mu + \sum_{i=1}^{p-1} \rho_i \Delta y_{t-i} + e_t$ and use standard t statistic to test $H_0: \mu = 0$
 - if H_0 can not be rejected continue to Step 5
 - if H_0 is rejected conclude that y_t is random walk with drift

Step 5. estimate model A and use τ_1 statistic to test H_0 : $\gamma=0$

- if H_0 can not be rejected conclude that y_t is random walk
- if H_0 is rejected conclude that y_t is trend stationary

```
library(urca)
ur.df(yTS, type = "trend", selectlags = "AIC") %>% summary()
```

```
##
## # Augmented Dickey-Fuller Test Unit Root Test #
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
## Residuals:
      Min
             10 Median
## -3.6246 -0.6734 -0.0073 0.6816 4.3585
## Coefficients:
              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2156769 0.0294252 7.330 2.68e-13 ***
## z.lag.1 -0.0562692 0.0047070 -11.954 < 2e-16 ***
## tt
            0.0084263 0.0007048 11.955 < 2e-16 ***
## z.diff.lag 0.0119032 0.0141433 0.842
## ---
## Signif, codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.018 on 4994 degrees of freedom
## Multiple R-squared: 0.02808. Adjusted R-squared: 0.02749
## F-statistic: 48.09 on 3 and 4994 DF. p-value: < 2.2e-16
##
## Value of test-statistic is: -11.9543 83.6306 71.4597
## Critical values for test statistics:
       1pct 5pct 10pct
## tau3 -3.96 -3.41 -3.12
## phi2 6.09 4.68 4.03
## phi3 8.27 6.25 5.34
```

ur.df(yTS[1:150], type = "trend", selectlags = "AIC") %>% summary()

```
##
** ***********************************
## # Augmented Dickey-Fuller Test Unit Root Test #
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
       Min
                10 Median
                                        Max
## -2.70057 -0.67726 -0.06942 0.71670 2.36169
## Coefficients:
              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.657770 0.284392
                                 2.313 0.0221 *
## z.lag.1 -0.088331 0.035947 -2.457 0.0152 *
## 1.1.
             0.009033 0.004035
                                 2.239 0.0267 *
## z.diff.lag -0.039590 0.082503 -0.480 0.6320
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.003 on 144 degrees of freedom
## Multiple R-squared: 0.04721. Adjusted R-squared: 0.02736
## F-statistic: 2.378 on 3 and 144 DF. p-value: 0.0723
##
## Value of test-statistic is: -2.4573 2.6964 3.0334
## Critical values for test statistics:
        1pct 5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2 6.22 4.75 4.07
## phi3 8.43 6.49 5.47
```

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test

- ▶ stationarity test H_0 : $\{y_t\}$ is stationary (either mean stationary or trend stationary) H_A : $\{y_t\}$ is difference stationary (has a unit root)
- \blacktriangleright main idea: decompose time series $\{y_t\}$ as

$$y_t = d_t + z_t + \varepsilon_t$$

where d_t is the deterministic trend, z_t is random walk $z_t=z_{t-1}+\nu_t$, ν_t is white noise (iid $E(\nu_t)=0$, $var(\nu_t)=\sigma_{\nu}^2$), and ε_t stationary error (i.e. I(0) but not necessarily white noise)

• stationarity of $\{y_t\}$ depends on σ_{ν}^2 , we can run a test

$$H_0: \sigma_{\nu}^2 = 0$$

against

$$H_A: \sigma_{\nu}^2 > 0$$

using Lagrange multiplier (LM) statistic

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test

▶ to perform KPSS test we estimate

model A
$$y_t = \mu + e_t$$

model B $y_t = \mu + \beta t + e_t$

model A is used if H_0 is mean stationarity, model B is used if H_0 is trend stationarity

lacktriangle using residuals e_t we construct LM statistics η

$$\eta = \frac{1}{T^2} \frac{1}{s^2} \sum_{t=1}^{T} S_t^2$$

where $S_t = \sum_{i=1}^t e_i$ is the partial sum process of the residuals e_t and s^2 is an estimator of the long-run variance of the residuals e_t .

▶ KPSS test is a one-sided right-tailed test: we reject H_0 at $\alpha\%$ level if η is greater than $100(1-\alpha)\%$ percentile from the appropriate asymptotic distribution

```
ur.kpss(yTS, type = "tau", lags = "long") %>% summary()
##
## # KPSS Unit Root Test #
** ****************
##
## Test is of type: tau with 31 lags.
##
## Value of test-statistic is: 0.1483
##
## Critical value for a significance level of:
                 10pct 5pct 2.5pct 1pct
## critical values 0.119 0.146 0.176 0.216
ur.kpss(vTS[1:150], type = "tau", lags = "long") %>% summary()
##
## # KPSS Unit Boot Test #
##
## Test is of type: tau with 13 lags.
##
## Value of test-statistic is: 0.1809
##
## Critical value for a significance level of:
                  10pct 5pct 2.5pct 1pct
##
## critical values 0.119 0.146 0.176 0.216
```

```
ur.kpss(yDS, type = "tau", lags = "long") %>% summary()
##
## # KPSS Unit Root Test #
** ****************
##
## Test is of type: tau with 31 lags.
##
## Value of test-statistic is: 1.9601
##
## Critical value for a significance level of:
                 10pct 5pct 2.5pct 1pct
## critical values 0.119 0.146 0.176 0.216
ur.kpss(vDS[1:150], type = "tau", lags = "long") %>% summary()
##
## # KPSS Unit Boot Test #
##
## Test is of type: tau with 13 lags.
##
## Value of test-statistic is: 0.1412
##
## Critical value for a significance level of:
                  10pct 5pct 2.5pct 1pct
##
## critical values 0.119 0.146 0.176 0.216
```

Phillips-Perron (PP) test

▶ an alternative to ADF test, estimates one of the models

$$\begin{array}{ll} \text{model A} & \Delta y_t = \gamma y_{t-1} + e_t \\ \\ \text{model B} & \Delta y_t = \gamma y_{t-1} + \mu + e_t \\ \\ \text{model C} & \Delta y_t = \gamma y_{t-1} + \mu + \beta t + e_t \end{array}$$

and tests
$$H_0: \gamma = 0$$
 against $H_A: \gamma < 0$

- ightharpoonup unlike ADF uses non-parametric correction based on Newey-West heteroskedasticity and autocorrelation consistent (HAC) estimators to account for possible autocorrelation in e_t
- advantage over the ADF: PP tests are robust to general forms of heteroskedasticity and do not require to choose number of lags in the test regression
- asymptotically identical to ADF test, but likely inferior in small samples
- like ADF also not very powerful at distinguishing stationary near unit root series for unit root series

Elliot, Rothenberg and Stock (ERS) tests

- two efficient unit root tests with substantially higher power than the ADF or PP tests especially when ϕ_1 is close to 1
- ▶ P-test: optimal for point alternative $\phi_1 = 1 \bar{c}/T$
- ▶ DF-GLS test: main idea estimate test regression as in model A of ADF but with detrended time series y_t

```
## ## ## Elliot, Rothenberg and Stock Unit Root Test #
## # Elliot, Rothenberg and Stock Unit Root Test #
## ## Elliot, Rothenberg and Stock Unit Root Test #
## ## Test of type P-test
## detrending of series with intercept and trend
##
## Critical values of P-test are:
## Critical values of P-test are:
## critical values 4.05 5.66 6.86
```

ur.ers(yTS, type="DF-GLS", model="trend") %>% summary()

```
##
## # Elliot, Rothenberg and Stock Unit Root Test #
## Test of type DF-GLS
## detrending of series with intercept and trend
##
## Call:
## lm(formula = dfgls.form, data = data.dfgls)
##
## Residuals:
     Min
             10 Median
## -3.5735 -0.7132 -0.0517 0.6432 4.2731
## Coefficients:
              Estimate Std. Error t value Pr(>|t|)
## vd.lag
            -0.041303 0.004285 -9.639 < 2e-16 ***
## yd.diff.lag1 0.003327 0.014217 0.234 0.81498
## vd.diff.lag2 -0.013141 0.014169 -0.927 0.35374
## vd.diff.lag3 -0.040292 0.014149 -2.848 0.00442 **
## vd.diff.lag4 0.002834 0.014147 0.200 0.84125
## ---
## Signif, codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 1.02 on 4990 degrees of freedom
## Multiple R-squared: 0.02337. Adjusted R-squared: 0.02239
## F-statistic: 23.88 on 5 and 4990 DF. p-value: < 2.2e-16
##
## Value of test-statistic is: -9.6387
## Critical values of DF-GLS are:
                 1pct 5pct 10pct
## critical values -3.48 -2.89 -2.57
```

ur.ers(yTS[1:150], type="DF-GLS", model="trend") %>% summary()

```
##
## # Elliot, Rothenberg and Stock Unit Root Test #
## Test of type DF-GLS
## detrending of series with intercept and trend
##
## Call:
## lm(formula = dfgls.form, data = data.dfgls)
##
## Residuals:
      Min
               10 Median
                                      Max
## -2.56982 -0.65834 -0.03218 0.73765 2.39730
## Coefficients:
              Estimate Std. Error t value Pr(>|t|)
## vd.lag
            -0.082652 0.036050 -2.293 0.0234 *
## vd.diff.lag1 -0.027003 0.084611 -0.319 0.7501
## vd.diff.lag2 -0.004045 0.083743 -0.048 0.9615
## vd.diff.lag3 -0.055587 0.083414 -0.666 0.5063
## vd.diff.lag4 0.092734 0.082401 1.125 0.2623
## ---
## Signif, codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.9947 on 140 degrees of freedom
## Multiple R-squared: 0.05753. Adjusted R-squared: 0.02387
## F-statistic: 1.709 on 5 and 140 DF, p-value: 0.1364
##
## Value of test-statistic is: -2.2927
## Critical values of DF-GLS are:
                 1pct 5pct 10pct
## critical values -3.46 -2.93 -2.64
```