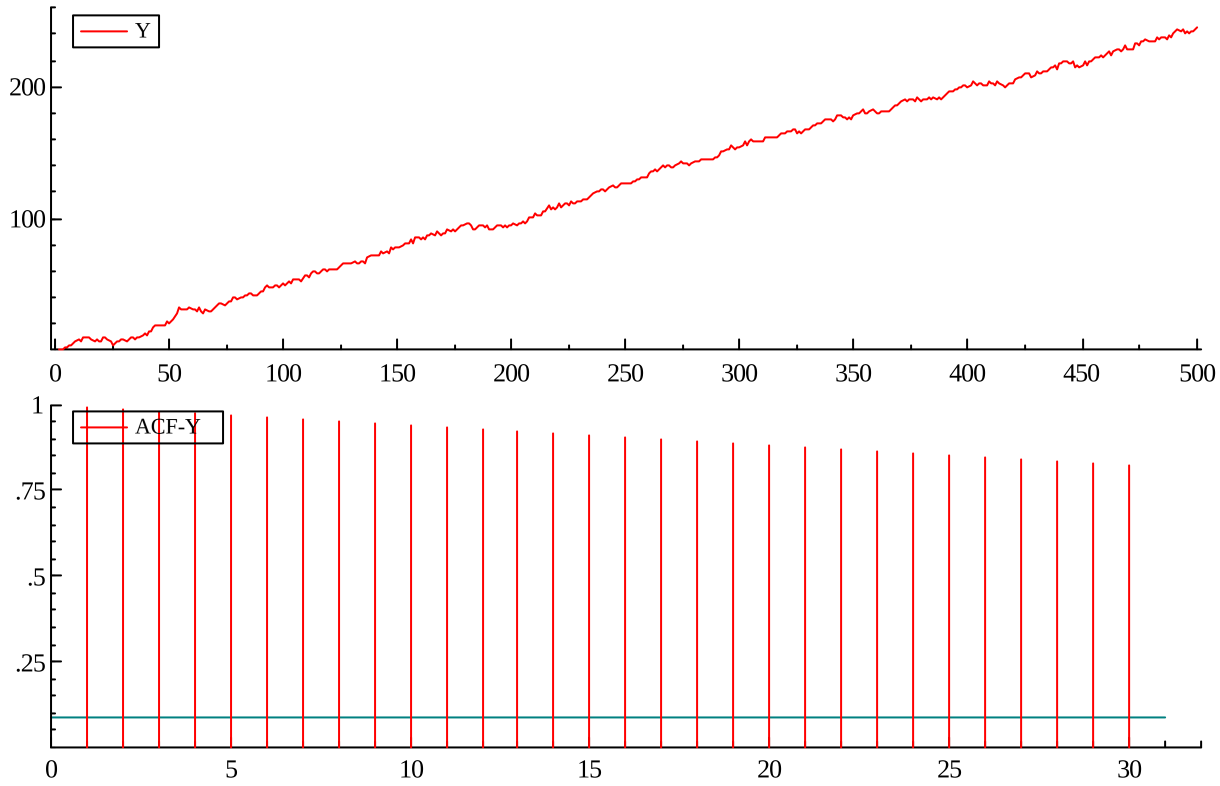
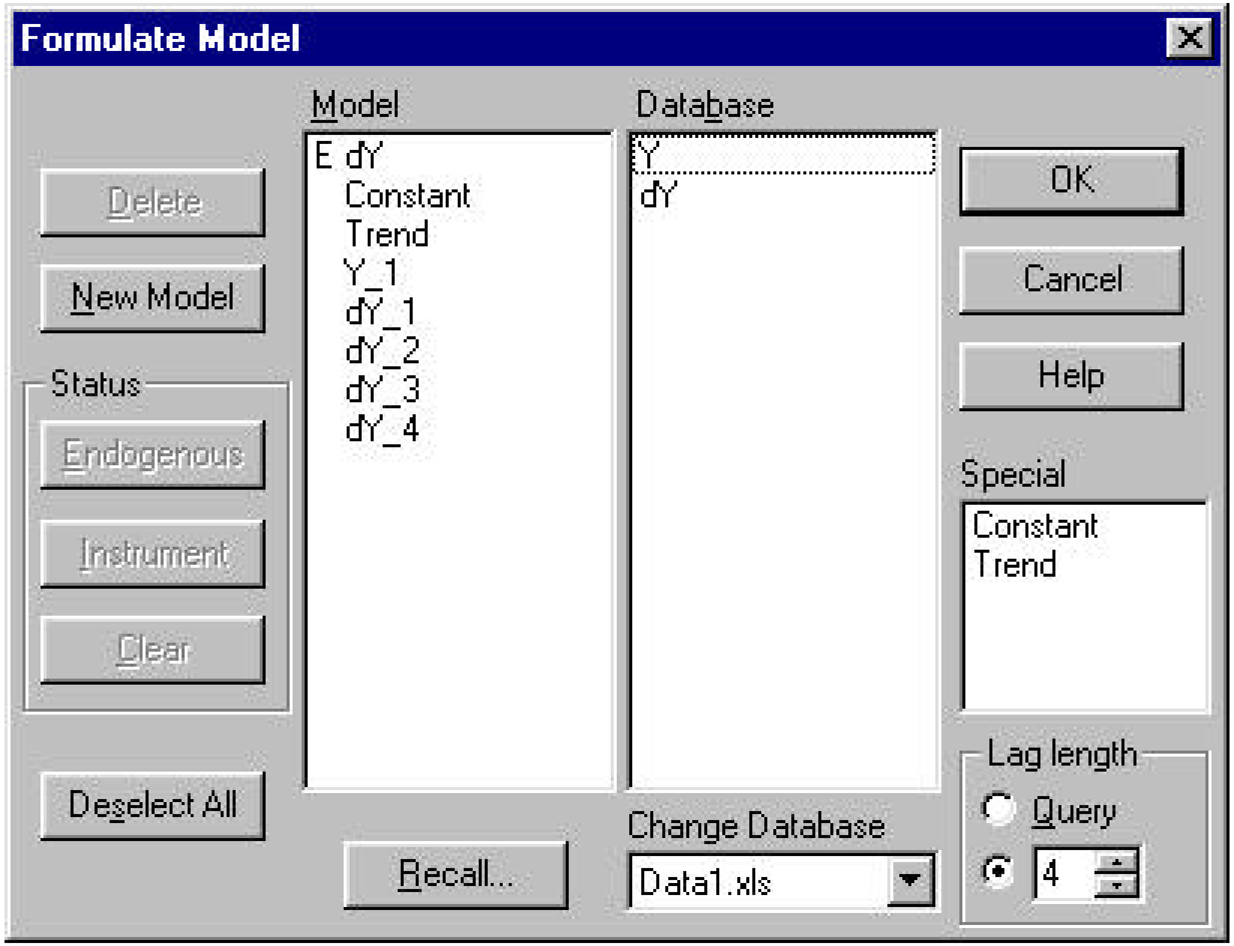
Unit Root Test

A Step By Step Approach

Consider The Following Series and Its Correlogram



This variable Y is clearly trended and you have to determine if this trend is stochastic or deterministic. After having created the difference variable ΔY estimate the following model, with as many lag of ΔY as you think appropriate.
(in the example I choose 4 lags of the variable ΔY)


Formulate Model
Model
Database
E dY
Constant
Trend
Y_1
dY_1
dY_2
dY_3
dY_4
Y
dY
Delete
New Model
Status
Endogenous
Instrument
Clear
Deselect All
Recall...
Change Database
Data1.xls
OK
Cancel
Help
Special
Constant
Trend
Lag length
Query
4

GiveWin - [Results]

File Edit View Tools Modules Window Help

```
EQ( 9) Modelling dY by OLS  (using Data1.xls)
The present sample is:  7 to 500

Variable      Coefficient    Std.Error  t-value  t-prob PartR^2
Constant          0.77619      0.14282    5.435  0.0000  0.0572
Trend            0.021665    0.0083743    2.587  0.0100  0.0136
Y_1             -0.044109     0.016875   -2.614  0.0092  0.0138
dY_1             -0.27120     0.046626   -5.817  0.0000  0.0650
dY_2            -0.059318     0.047966   -1.237  0.2168  0.0031
dY_3            -0.068276     0.047749   -1.430  0.1534  0.0042
dY_4            -0.043802     0.045617   -0.960  0.3374  0.0019

R^2 = 0.0996592  F(6,487) = 8.9844 [0.0000]  \sigma = 1.40494  DW = 1.99
RSS = 961.2700903 for 7 variables and 494 observations


AR 1- 2 F( 2,485) =   0.033719 [0.9668]
ARCH 1  F( 1,485) =     3.0964 [0.0791]
Normality Chi^2(2)=     0.4636 [0.7931]
Xi^2    F(12,474) =    0.99731 [0.4503]
Xi*Xj   F(27,459) =    0.95309 [0.5346]
RESET   F( 1,486) =   0.029887 [0.8628]
```

Ready | Data1.xls | PcGive | Line 25

```
EQ(10) Modelling dY by OLS  (using Data1.xls)
The present sample is:   7 to 500

Variable     Coefficient    Std.Error  t-value  t-prob PartR^2
Constant        0.74235     0.13839     5.364  0.0000  0.0557
Trend          0.022755    0.0082963    2.743  0.0063  0.0152
Y_1           -0.046289    0.016721    -2.768  0.0058  0.0155
dY_1           -0.26688    0.046405    -5.751  0.0000  0.0635
dY_2          -0.054697    0.047720    -1.146  0.2523  0.0027
dY_3          -0.054428    0.045516    -1.196  0.2324  0.0029

R^2 = 0.0979547  F(5,488) = 10.599 [0.0000]  \sigma = 1.40483  DW = 1.99
RSS = 963.0900147 for 6 variables and 494 observations


AR 1- 2 F( 2,486) =     0.50657  [0.6029]
ARCH 1  F( 1,486) =      2.6624  [0.1034]
Normality Chi^2(2)=     0.52037  [0.7709]
Xi^2     F(10,477) =     1.0716  [0.3825]
Xi*Xj    F(20,467) =    0.92018  [0.5614]
RESET    F( 1,487) =   0.017673  [0.8943]
```

```
EQ(11) Modelling dY by OLS  (using Data1.xls)
The present sample is:   7 to 500

Variable     Coefficient    Std.Error  t-value  t-prob PartR^2
Constant        0.70291     0.13446     5.228  0.0000  0.0529
Trend          0.024205    0.0082109    2.948  0.0034  0.0175
Y_1           -0.049186    0.016551    -2.972  0.0031  0.0177
dY_1           -0.26185    0.046234    -5.664  0.0000  0.0616
dY_2          -0.037629    0.045555    -0.826  0.4092  0.0014

R^2 = 0.0953115  F(4,489) = 12.879 [0.0000]  \sigma = 1.40545  DW = 1.99
RSS = 965.9120909 for 5 variables and 494 observations


AR 1- 2 F( 2,487) =     0.82392  [0.4393]
ARCH 1  F( 1,487) =      2.4217  [0.1203]
Normality Chi^2(2)=     0.46115  [0.7941]
Xi^2     F( 8,480) =     1.2016  [0.2960]
Xi*Xj    F(14,474) =     1.0893  [0.3647]
RESET    F( 1,488) =   0.028713  [0.8655]
```

```
EQ(12) Modelling dY by OLS  (using Data1.xls)
The present sample is:   7 to 500

Variable     Coefficient    Std.Error  t-value  t-prob PartR^2
Constant        0.67820     0.13105     5.175  0.0000  0.0518
Trend          0.025237    0.0081125    3.111  0.0020  0.0194
Y_1           -0.051254    0.016356    -3.134  0.0018  0.0196
dY_1           -0.25042    0.044100    -5.678  0.0000  0.0617

R^2 = 0.0940491  F(3,490) = 16.956 [0.0000]  \sigma = 1.40499  DW = 2.01
RSS = 967.259836 for 4 variables and 494 observations


AR 1- 2 F( 2,488) =     0.95032  [0.3873]
ARCH 1  F( 1,488) =      1.9438  [0.1639]
Normality Chi^2(2)=     0.22312  [0.8944]
Xi^2     F( 6,483) =     1.4386  [0.1979]
Xi*Xj    F( 9,480) =     1.4197  [0.1766]
RESET    F( 1,489) = 0.00076464  [0.9780]
```

```
EQ(13) Modelling dY by OLS  (using Data1.xls)
The present sample is:   7 to 500

Variable     Coefficient    Std.Error  t-value  t-prob PartR^2
Constant        0.55581     0.13331     4.169  0.0000  0.0342
Trend          0.034153    0.0082085    4.161  0.0000  0.0341
Y_1           -0.069168    0.016551    -4.179  0.0000  0.0343

R^2 = 0.0344316  F(2,491) = 8.7544 [0.0002]  \sigma = 1.44901  DW = 2.45
RSS = 1030.911951 for 3 variables and 494 observations


AR 1- 2 F( 2,489) =      16.468  [0.0000] **
ARCH 1  F( 1,489) =       4.417  [0.0361] *
Normality Chi^2(2)=    0.022467  [0.9888]
Xi^2     F( 4,486) =      1.6097  [0.1706]
Xi*Xj    F( 5,485) =      1.4564  [0.2027]
RESET    F( 1,490) =     0.18765  [0.6651]
```

Choose Between Alternative Models - The Model-Progress Results

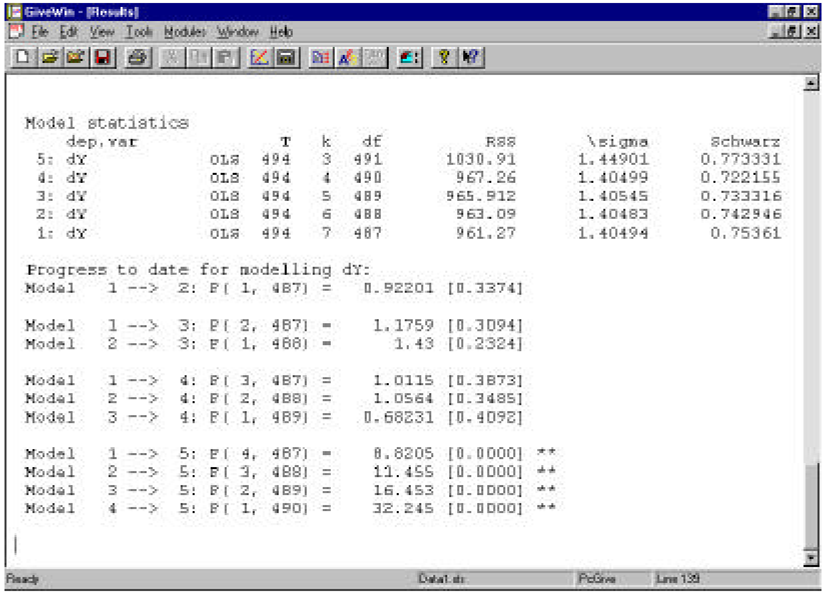
Model 1: $\Delta Y_t = b_0 + b_1 t + bY_{t-1} + a_1 \Delta Y_{t-1} + a_2 \Delta Y_{t-2} + a_3 \Delta Y_{t-3} + a_4 \Delta Y_{t-4} + e$

Model 2: $\Delta Y_t = b_0 + b_2 t + bY_{t-1} + a_1 \Delta Y_{t-1} + a_2 \Delta Y_{t-2} + a_3 \Delta Y_{t-3} + e$

Model 3: $\Delta Y_t = b_0 + b_2 t + bY_{t-1} + a_1 \Delta Y_{t-1} + a_2 \Delta Y_{t-2} + e$

Model 4: $\Delta Y_t = b_0 + b_2 t + bY_{t-1} + a_1 \Delta Y_{t-1} + e$

Model 5: $\Delta Y_t = b_0 + b_2 t + bY_{t-1} + e$



```
Model statistics
    dep.var                 T    k   df          RSS      \sigma     Schwarz
 5: dY              OLS   494    3  491      1030.91     1.44901    0.773331
 4: dY              OLS   494    4  490       967.26     1.40499    0.722155
 3: dY              OLS   494    5  489      965.912     1.40545    0.733316
 2: dY              OLS   494    6  488       963.09     1.40483    0.742946
 1: dY              OLS   494    7  487       961.27     1.40494     0.75361

Progress to date for modelling dY:
Model   1 -->  2: F( 1, 487) =   0.92201 [0.3374]

Model   1 -->  3: F( 2, 487) =    1.1759 [0.3094]
Model   2 -->  3: F( 1, 488) =      1.43 [0.2324]

Model   1 -->  4: F( 3, 487) =    1.0115 [0.3873]
Model   2 -->  4: F( 2, 488) =    1.0564 [0.3485]
Model   3 -->  4: F( 1, 489) =   0.68231 [0.4092]

Model   1 -->  5: F( 4, 487) =    8.8205 [0.0000] **
Model   2 -->  5: F( 3, 488) =    11.455 [0.0000] **
Model   3 -->  5: F( 2, 489) =    16.453 [0.0000] **
Model   4 -->  5: F( 1, 490) =    32.245 [0.0000] **
```

Both the F-Test and the Schwarz Information Criteria indicates that MODEL 4 is the one to be preferred

Unit Root Testing

After having estimated, according to the previous analysis, the following equation

$$\Delta Y_t = b_0 + b_2 t + \boldsymbol{b} Y_{t-1} + \boldsymbol{a}_1 \Delta Y_{t-1} + \boldsymbol{e}$$

The alternative hypothesis are (in this particular case)

$$H_0 : (b_0, b_2, \boldsymbol{b}) = (b_0, 0, 0)$$

$$v$$

$$H_1 : (b_0, b_2, \boldsymbol{b}) \neq (b_0, 0, 0)$$

To do this perform a F-Test and use the

$$\boldsymbol{f}_3$$

Result

```
Wald test for linear restrictions: Subset
LinRes  F( 2,493) =     5.0781 [0.0066] **
```

Therefore we can reject the hypothesis that we have a trend in the equation. We can continue the analysis using

$$\Delta Y_t = b_0 + bY_{t-1} + a_1 \Delta Y_{t-1} + e$$

Use

f_1 Statistics - Use the F statistic to check if $\mathbf{b} = \mathbf{b_0} = \mathbf{0}$ using the non standard tables

t_m Statistics - use the t statistics to check if $\mathbf{b}=\mathbf{0}$, again using non-standard tables

```
EQ( 2) Modelling dY by OLS  (using Data1.xls)
The present sample is:  4 to 500

Variable      Coefficient    Std.Error  t-value  t-prob PartR^2
Constant          0.69327      0.13033    5.319  0.0000  0.0542
dY_1             -0.27787     0.043463   -6.393  0.0000  0.0764
Y_1           -0.00051485   0.00089136   -0.578  0.5638  0.0007

R^2 = 0.0769002  F(2,494) = 20.577 [0.0000]  \sigma = 1.41413  DW = 2.02
RSS = 987.8846244 for 3 variables and 497 observations


AR 1- 2 F( 2,492) =     2.3776 [0.0938]
ARCH 1  F( 1,492) =     2.0004 [0.1579]
Normality Chi^2(2)=    0.18218 [0.9129]
Xi^2    F( 4,489) =     1.7354 [0.1409]
Xi*Xj   F( 5,488) =     1.7202 [0.1283]
RESET   F( 1,493) =   0.001058 [0.9741]

Wald test for linear restrictions: Subset
LinRes  F( 2,494) =     44.403 [0.0000] **

Zero restrictions on:
 Constant Y_1
```

The t-stat cannot reject the hypothesis of Unit Root while the F-stat reject the hypothesis that the drift is equal to zero. Therefore we can conclude that the model most likely to describe the true DGP is

$$\Delta Y_t = b_0 + Y_{t-1} + a_1 \Delta Y_{t-1} + e$$

Look at the Series - There is a Trend?

- **Yes**
  - Estimate
    $$\Delta Y_t = b_0 + b_2 t + bY_{t-1} + \sum a_j \Delta Y_{t-j} + e$$
  - Use f_3 to test
    $$H_0: (b_0, b_2, b) = (b_0, 0, 0)$$
    $$v$$
    $$H_1: (b_0, b_2, b) \neq (b_0, 0, 0)$$
    - Reject
      - test β=0 using the t-stat. from step 1 using t_t
        - Reject
          - **No Unit Root**
            - Normal Test procedure to determine the presence of Time trend or Drift
        - Accept
          - **Unit Root +Trend**
            - Use f_2
              To determine if there is a drift as well
    - Accept → (go to the "No" branch estimate)

- **No**
  - Estimate
    $$\Delta Y_t = b_0 + bY_{t-1} + \sum a_j \Delta Y_{t-j} + e$$
  - Use f_1 to test
    $$H_0: (b_0, b) = (0,0)$$
    $$v$$
    $$H_1: (b_0, b) \neq (0,0)$$
    - Reject
      - test β=0 using the t-stat. from step 1 using t_m
        - Reject
          - Stable Series, use normal test to check the drift
        - Accept
          - Random Walk + Drift
    - Accept
      - Pure Random Walk