Lecture 3: Regression analysis with time series data II

Prof. dr. Wolter Hassink
Utrecht University School of Economics
w.h.j.hassink@uu.nl

Contents:

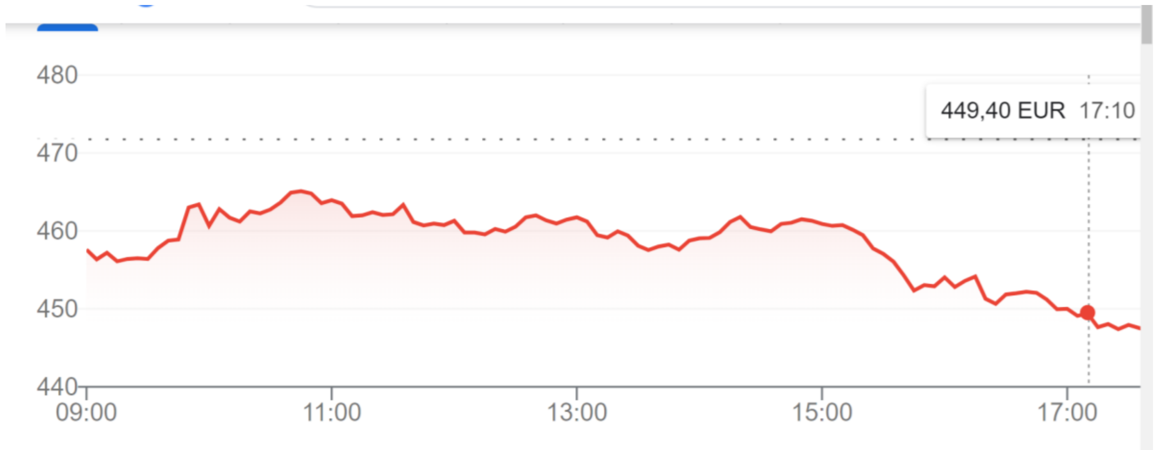
- Spurious regression
- Consequences of spurious regression & stationarity
- Dickey-Fuller test
- Application of Dickey-Fuller test
- Co-integration
- Error-correction models
- Forecasting
- Vector autoregressive (VAR) model

Material:

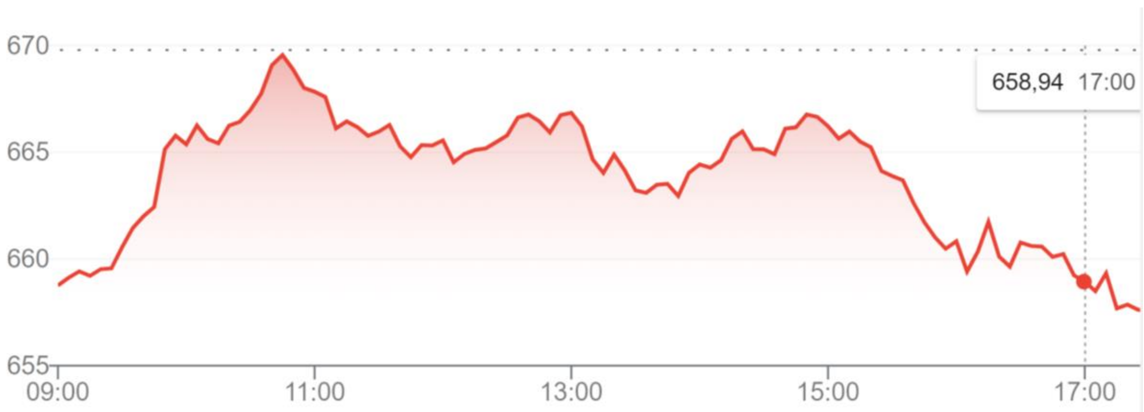
Chapter 18: 11.3, 18.2, 18.3, 18.4, 18.5

Motivation

- For instance the stockprice of ASML. Development of ASML on Thursday 22 September 2023



- There is also information on the development of the AEX index on the same day



FORECASTING THE STOCKPRICE OF ASML – WAY OF THINKING

- Construct a forecasting model. Make a sample split into the observations to estimate the parameters of a forecasting model
- The remainder of the observations are used to make predictions (forecasts) with the forecasting model
- Make use of the model that has the smallest forecast error for the testing dataset (= running a horse race)
- Options:
    1. simple time trend: $AS\hat{M}L_t = \hat{\alpha}_0 + \hat{\alpha}_1 t$
    2. Lagged values: $AS\hat{M}L_t = \hat{\beta}_0 + \hat{\beta}_1 ASML_{t-1}$ with $\left|\hat{\beta}_1\right| < 1$
    3. Other variables: $AS\hat{M}L_t = \hat{\gamma}_0 + \hat{\gamma} AEX_{t-1}$
    4. Unit root/Random walk/I(1): $AS\hat{M}L_t = \hat{\lambda}_0 + ASML_{t-1}$
    5. Co-integration: $AS\hat{M}L_t = \hat{\delta}_0 + \hat{\delta}_1 AEX_t$
    6. Some combination of these five options
    7. Vector Autoregressive (VAR) model

- Previous week: options 1, 2 and 3
- THIS WEEK: options 4 and 5 and 7
- THIS WEEK: shall we analyse $ASML_t$ or $\Delta ASML_t \quad (= ASML_t - ASML_{t-1})$

  We will apply the Dickey-Fuller test

Spurious regression

Motivation: spurious regression

Aim: to introduce spurious regression and co-integration.

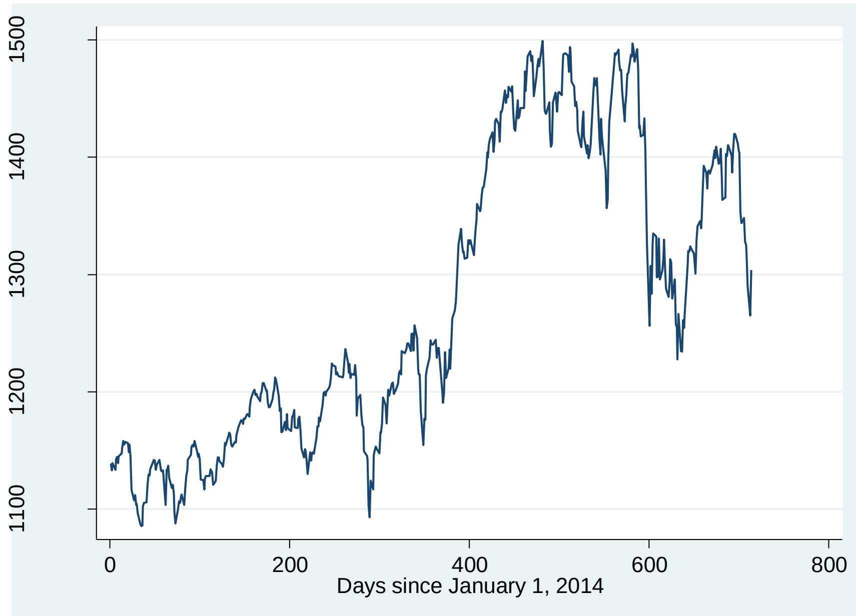
- Consider a regression of y_t on x_t where both variables are non-stationary (they are integrated of order one – I(1)). A good example is the relation between AEXindex and the stockprice of ASML.
- In that case, there is the possibility of spurious regression. CONSEQUENCE: Estimation by OLS gives inconsistent coefficients and the *t*-values are misleading.
- Spurious regression does not occur if the two variables are co-integrated, that is, a linear combination of the two variables is stationary (I(0)).

  Example - Spurious regression
- Let y_t (= stock price ASML) and x_t (=AEX index) be two independent random walks:
  - $y_t = y_{t-1} + e_t$
  - $x_t = x_{t-1} + a_t$

  Where the error terms e_t and a_t are mutually independent.
- Let's assume that y_t and x_t are independent, so that there is no relationship between both variables.
- However, a regression of the level of *y* (not a growth variable) on the level of *x* yields a high R^2 and a very low Durbin Watson statistic:
  - $y_t = \beta_0 + \beta_1 x_t + u_t$ (18.30)
- Adding a time trend to the regression equation does not change the result, and the results cannot be trusted.
  - $y_t = \beta_0 + \beta_1 x_t + \gamma t + u_t$
- However, a regression in first differences delivers the expected results: $\Delta y_t = \alpha_2 + \beta_2 \Delta x_t + \Delta u_t$

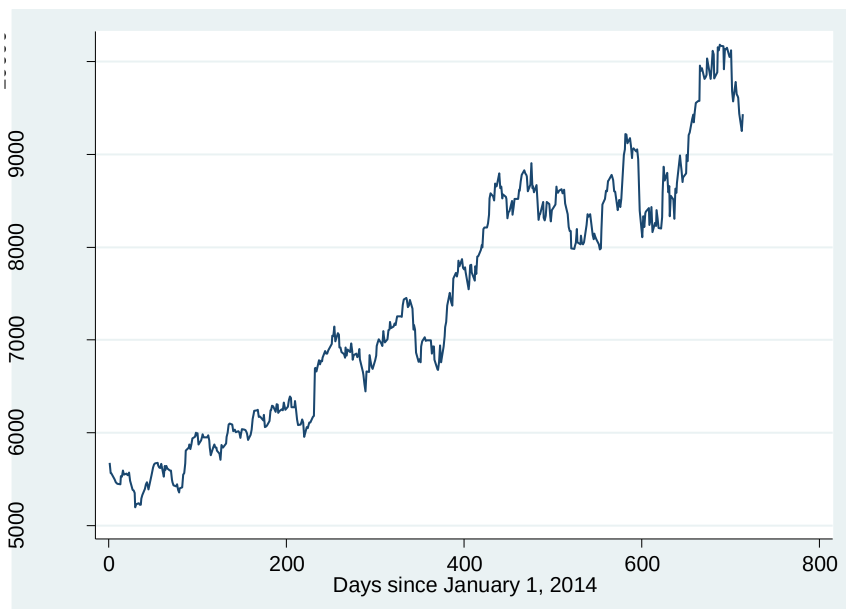
Level: *AEXindex*

```
. keep if stock == 14
. tsset time
. graph twoway line AEXindex time
```



Level: *Stockprice*

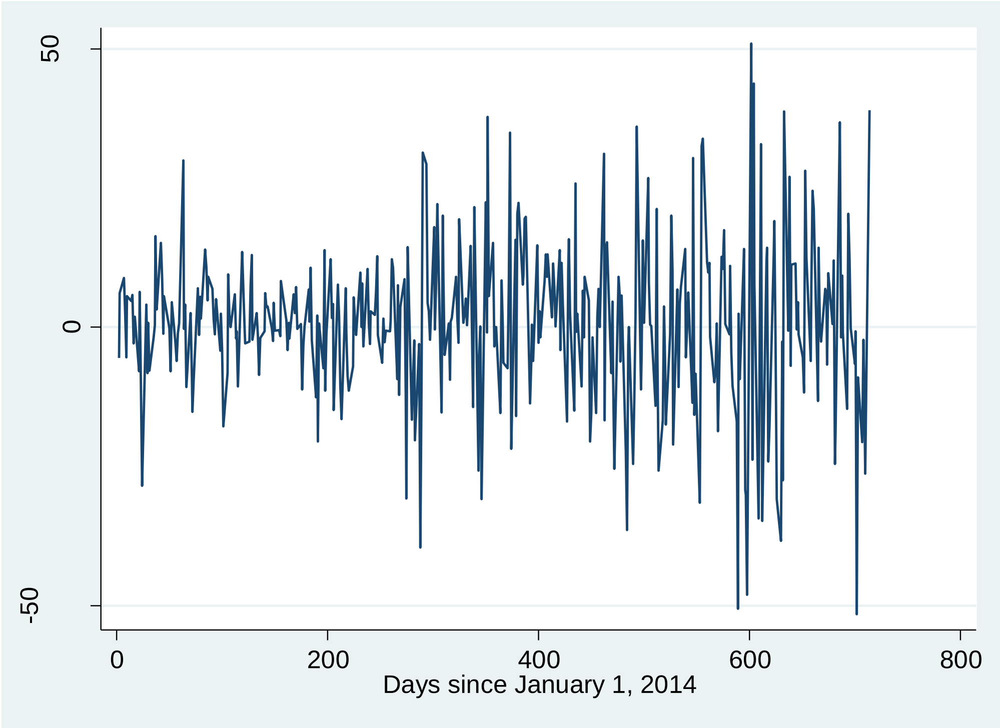
```
. graph twoway line stockprice time
```



The purpose of those graphs is to recognize the development of the level stockprice and to check whether any time trend is needed. Graphs below: the development in the first difference of the stock price is not completely irregular. This is what will be usually checked first.

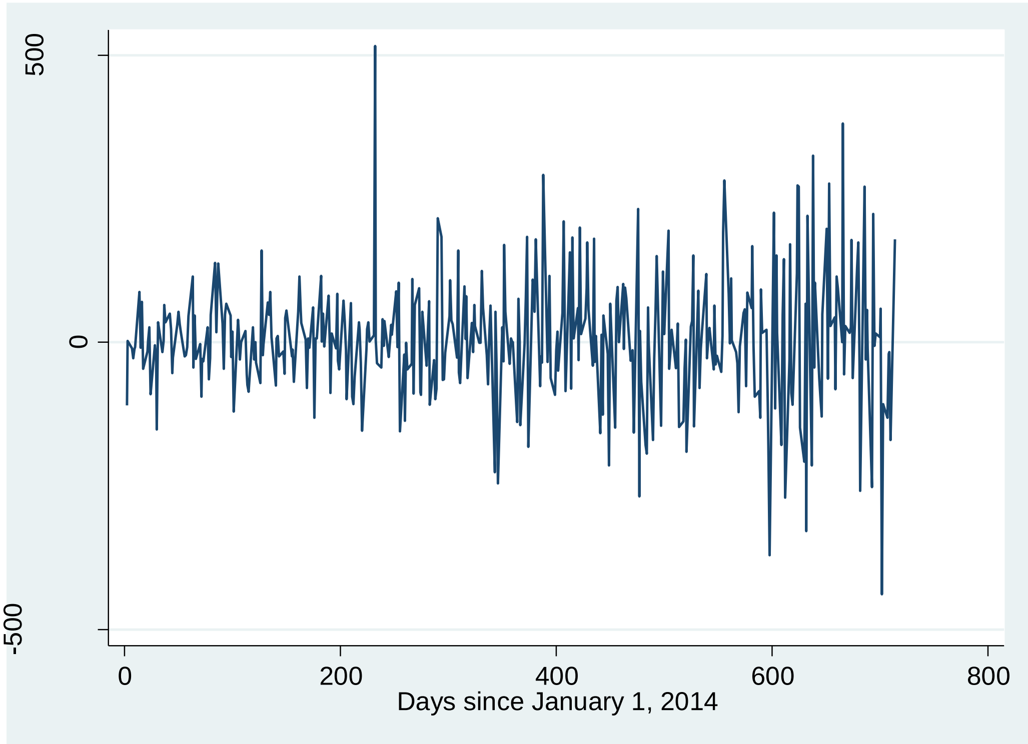
First difference: $\Delta AEXindex$

```
. graph twoway line d.AEXindex time
```



First difference: $\Delta stockprice$

```
. graph twoway line d.stockprice time
```



Consequences of spurious regression & stationarity

Time series: consistency of OLS

Consider a time-series model with a finite lag.

$$y_t = \alpha + \delta_0 z_t + \delta_1 z_{t-1} + u_t \qquad t = 1, 2, ...$$

- The regression parameters $\alpha, \delta_0, \delta_1$ can be estimated consistently with OLS if the following assumptions hold (in addition to the assumptions of linearity and no perfect collinearity):

TS.1' y_t, z_t, z_{t-1}, u_t must be:

- Stationary
- Weakly dependent.

TS.3' z_t, z_{t-1} are contemporaneously exogenous: $Cov(z_t, u_t) = 0$ and $Cov(z_{t-1}, u_t) = 0$.

Week 2: weak dependence

- Moving average (MA) and autoregressive (AR) models are weakly dependent, in contrast to unit root models.

Correlation between x_t and x_{t+1}:

$$Corr(x_t, x_{t+1}) = \frac{Cov(x_t, x_{t+1})}{\sqrt{Var(x_t)}\sqrt{Var(x_{t+1})}}$$

Correlation between x_t and x_{t+2}:

$$Corr(x_t, x_{t+2}) = \frac{Cov(x_t, x_{t+2})}{\sqrt{Var(x_t)}\sqrt{Var(x_{t+2})}}$$

……….

Correlation between x_t and x_{t+h}:

$$Corr(x_t, x_{t+h}) = \frac{Cov(x_t, x_{t+h})}{\sqrt{Var(x_t)}\sqrt{Var(x_{t+h})}}$$

The definition of **weakly-dependent time series:**

A stationary time series $\{x_t : t = 1, 2, ...,\}$ is weakly dependent if x_t and x_{t+h} are "almost independent" as $h \to \infty$. Thus $Corr(x_t, x_{t+h}) \to 0$ as $h \to \infty$.

Stationarity

Formal definition of stationary stochastic process (see Wooldridge, Section 11.1):
The stochastic process $\{x_t : t = 1, 2, ...,\}$ is stationary if for every collection of time indices $1 \leq t_1 < t_2 < ... < t_m$, the joint distribution of $x_{t_1} < x_{t_2} < ... < x_{t_m}$ is the same as the joint distribution of $x_{t_1+h} < x_{t_2+h} < ... < x_{t_m+h}$ for all integers $h \geq 1$.

- A stochastic process that is not stationary, is said to be a nonstationary process.
- Stationarity implies that the sequence $\{x_t : t = 1, 2, ...,\}$ is *identically distributed*.
- A process with a time trend is nonstationary

Weaker requirement: covariance stationarity

The stochastic process $\{x_t : t = 1, 2, ...,\}$ is covariance stationary if
1) $E(x_t)$ is constant
2) $Var(x_t)$ is constant
3) For any *t*, $h \geq 1$, $Cov(x_t, x_{t+h})$ depends only on *h* and not on *t*.

Examples of stationary and non-stationary models

Three classes of models are considered:
1) Moving Average (MA) models. E.g. MA(1) (maximum of one period lag):

$$y_t = e_t + \alpha_1 e_{t-1} \qquad t = 1, 2, ...$$

where the error term e_t is i.i.d. with $Ee_t = 0$ and $Var(e_t) = \sigma_e^2$
Features: y_t is stationary and weakly dependent.

2) Autoregressive models. E.g. AR(1) (maximum of one period lag):

$$y_t = \rho_1 y_{t-1} + e_t \qquad t = 1, 2, ...$$

where e_t is i.i.d. with $Ee_t = 0$ and $Var(e_t) = \sigma_e^2$. The error term e_t is independent of y_{t-1}.
Features: y_t is stationary and weakly dependent if the stability condition is met: $|\rho_1| < 1$.

3) Unit root model:

$$y_t = y_{t-1} + e_t \qquad t = 1, 2, ...$$

where e_t is i.i.d. with $Ee_t = 0$ and $Var(e_t) = \sigma_e^2$. The error term e_t is independent of y_{t-1}. Features: y_t is a highly persistent time series. It is nonstationary and not weakly dependent.

We would like to know how to respond to models containing unit root.

Response:
1) Do not need to apply any transformation to the time series if a time series does not have a unit root.
2) Take the first differences if a time series has a unit root.

- $\{y_t\}$: Integrated of order one (or I(1)).

$\{y_t\}$:I(1) or $y_t = y_{t-1} + e_t$

- Note that the first difference of $\{y_t\}$ is $\{e_t\}$.

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

Since e_t is assumed to be i.i.d., it is stationary and weakly dependent.

Dickey-Fuller test

How do we check for a unit root? Dickey-Fuller test

Aim: to describe the Dickey-Fuller(DF) test.

- Start with an AR(1) model:

$$y_t = \alpha + \rho y_{t-1} + e_t \qquad t = 1, 2, ..., n \qquad (18.17)$$

- Assume that e_t is i.i.d. with zero mean and constant variance. e_t is independent of $y_{t-1}, y_{t-2}, ..., y_0$.:

  $$E(e_t \mid y_{t-1}, y_{t-2}, ..., y_0) = 0$$

- $\rho = 1$ gives a random walk
  - $\alpha = 0$ without drift
  - $\alpha \neq 0$ with drift
- The hypotheses:
  - $H_0 : \rho = 1$ (unit root)

    $H_1 : \rho < 1$ (no unit root)
- Note:
  - Under the null hypothesis the time series $\{y_t\}$ follows random walk/$\{y_t\}$ has a unit root/ $\{y_t\}$ is I(1) (integrated of order one). These three terms are synonyms. The time series does not follow a *t*-distribution if the null hypothesis is not rejected. Not even asymptotically.
  - The alternative hypothesis is one sided. If the null hypothesis is rejected, the time series follows a stable AR(1) model.

- Define: $\theta = \rho - 1$ (18.19)
- $y_t = \alpha + \rho y_{t-1} + e_t$ becomes

$$y_t - y_{t-1} = \alpha + \rho y_{t-1} - y_{t-1} + e_t$$
$$\Delta y_t = \alpha + (\rho - 1) y_{t-1} + e_t \qquad (18.21)$$

- which is a regression of Δy_t on y_{t-1}:

- The hypotheses become:
  - $H_0 : \theta = 0; (\rho = 1)$ (unit root)
    $H_1 : \theta < 0; (\rho < 1)$ (no unit root)
- Note:
  - The left-hand side is the first difference variable Δy_t (which is I(0) if H_0 is true); the right-hand side is a level variable (I(1) if H_0 is true).
  - Problem: the *t*-statistic for $\hat{\theta}$ does not follow *t*-distribution under H_0 as y_{t-1} is I(1).
  - Dickey-Fuller uses different critical values for the *t*-statistic of $\hat{\theta}$ (estimated parameter on lagged level variable)

Table 18.2

Signif. level	1%	2.5%	5%	10%
Critical value	-3.43	-3.12	-2.86	-2.57

- E.g. reject H_0 if *t*-value < -2.86. Compared with the standard *t*-test critical values (e.g. –1.645 with α=0.05), it is more difficult to reject H_0 with the DF-critical values.
- Apply DF-test on all variables of the regression equation.

Augmented Dickey-Fuller test

Aim: to introduce the augmented DF-test, which corrects for autocorrelation in the residuals.

- Consider a model with additional lags

$$\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t \qquad \text{with } \gamma_1 < 1 \qquad (18.23)$$

$$t = 2, 3, \ldots, n$$

- DF-test: H_0: $\theta = 0$; H_1: $\theta < 0$ using DF-critical values of *t*-statistic $\hat{\theta}$
- The critical values of *t*-statistic for $\hat{\gamma}_1$ (lagged dependent variable) approximately follow a *t*-distribution. Standard *t*-tests and *F*-tests can be applied to determine the appropriate number of (p) lags, Δy_{t-j} to include. $\quad j = 1, 2, \ldots, p$
- Rule of thumb:
  - Annual data: one or two lags suffice
  - Monthly data: 12 lags might be needed
  - Quarterly data: four or five lags.

Dickey-Fuller test with time trend

Aim: to introduce a time trend in the DF-test.

- A time trend is added to the basic equation:

$$\Delta y_t = \alpha + \delta t + \theta y_{t-1} + e_t \qquad t = 2,3,...,n \tag{18.25}$$

- The same testing procedure as above is applied. Problem: *t*-value of the estimated parameter on time trend $\hat{\delta}$ does not follow *t*-distribution under H_0.
- Dickey-Fuller: the critical values for the *t*-statistic of $\hat{\theta}$ (estimated parameter on lagged level variable) are:

Table 18.3

Signif. level	1%	2.5%	5%	10%
Critical value	-3.96	-3.66	-3.41	-3.12

- E.g. reject H_0 if *t*-value < -3.41 at the 5% level. The critical value (5% significance level) of Table 18.2 was 2.86.
- The problem with the DF-test is that it lacks power (large type II-error). There is a high tendency to “accept” the null hypothesis of a unit root, while H_1 (stationarity) is true. This is especially true in the case of small samples.

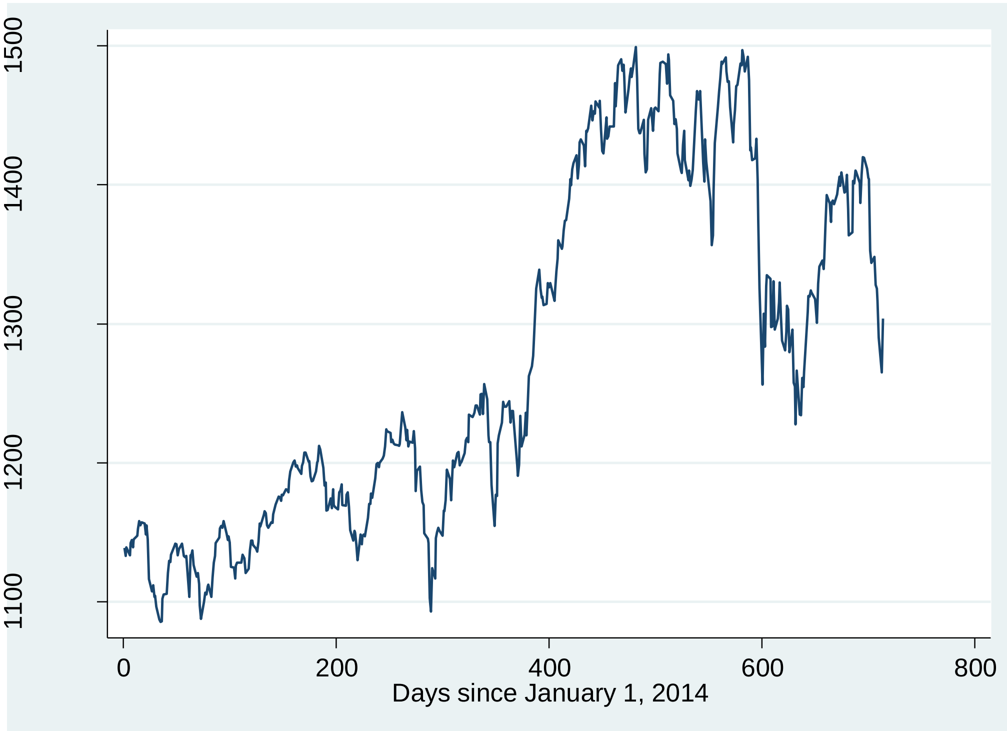
Application of Dickey-Fuller test

Example 1: stockdata.dta

Structure:

- Visually inspect to determine whether a time trend is needed in the test.
- Test for unit root in both *AEXindex* and *stockprice*.
- Start with augmented Dickey-Fuller test and determine the appropriate number of lags to include.
- Check whether the time trend is needed in the DF-test.

graph twoway line AEXindex time



1. Test for *stockprice* (no time trend)

- First investigate how many lags of the dependent variable should be included (augmented Dickey-Fuller test).

```
. reg d.AEXindex  l.d.AEXindex l2.d.AEXindex l3.d.AEXindex l.AEXindex

      Source |       SS       df       MS              Number of obs =     506
-------------+------------------------------           F(  4,   501) =    1.13
       Model |  962.017589     4  240.504397           Prob > F      =  0.3399
    Residual |  106309.236   501  212.194083           R-squared     =  0.0090
-------------+------------------------------           Adj R-squared =  0.0011
       Total |  107271.253   505  212.418323           Root MSE      =  14.567
------------------------------------------------------------------------------
  D.AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    AEXindex |
         LD. |   .0386363   .0448846     0.86   0.390    -.0495491    .1268216
        L2D. |  -.0332562   .0450354    -0.74   0.461    -.1217377    .0552252
        L3D. |   .0445622   .0451556     0.99   0.324    -.0441554    .1332798
         L1. |  -.0083304   .0052092    -1.60   0.110     -.018565    .0019042
             |
       _cons |   10.95294   6.677218     1.64   0.102    -2.165858    24.07174
------------------------------------------------------------------------------

. test  l.d.AEXindex l2.d.AEXindex l3.d.AEXindex

 ( 1)  LD.AEXindex = 0
 ( 2)  L2D.AEXindex = 0
 ( 3)  L3D.AEXindex = 0

       F(  3,   501) =    0.71
            Prob > F =    0.5479
```

$H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0$

$H_1 : H_0$ not true

The F-statistic > critical value then reject H_0

Since 0.71 < 2.60 we cannot reject H_0

This can also be observed from the p-value: 0.5479 > 0.05

```
. reg d.AEXindex  l.d.AEXindex  l.AEXindex

      Source |       SS       df       MS              Number of obs =     508
-------------+------------------------------           F(  2,   505) =    1.52
       Model |  641.309471     2  320.654735           Prob > F      =  0.2202
    Residual |  106703.314   505  211.293691           R-squared     =  0.0060
-------------+------------------------------           Adj R-squared =  0.0020
       Total |  107344.624   507  211.725096           Root MSE      =  14.536

------------------------------------------------------------------------------
  D.AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    AEXindex |
         LD. |   .0354617   .0447228     0.79   0.428     -.052404    .1233274
         L1. |  -.0082106   .0051726    -1.59   0.113    -.0183732    .0019519
             |
       _cons |   10.80051   6.629097     1.63   0.104    -2.223492    23.82452
------------------------------------------------------------------------------
```

- Estimated parameter on D.AEX_{t-1}: a one unit increase in the previous period´s change in the AEX index is associated with a 0.0355 unit increase in the current period´s change in the AEX index (the t-statistic of 0.79 indicates the effect is zero)

- Estimated parameter on AEX_{t-1}: a one unit increase in AEX in the previous period is associated with a -0.0082 unit decrease in the change in AEX in the current period. (the t-statistic of -1.59 indicates the effect is zero)

```
. reg d.AEXindex  l.AEXindex

      Source |       SS       df       MS              Number of obs =     509
-------------+------------------------------           F(  1,   507) =    2.35
       Model |   494.45124     1   494.45124           Prob > F      =  0.1263
    Residual |  106884.985   507   210.81851           R-squared     =  0.0046
-------------+------------------------------           Adj R-squared =  0.0026
       Total |  107379.436   508  211.376842           Root MSE      =   14.52

------------------------------------------------------------------------------
  D.AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    AEXindex |
         L1. |  -.0078945   .0051549    -1.53   0.126     -.018022     .002233
             |
       _cons |   10.39236   6.605534     1.57   0.116    -2.585224    23.36995
------------------------------------------------------------------------------
```

H_0: $\theta = 0$ (thus: $\rho = 1$)
H_1: $\theta < 0$ (thus: $\rho < 1$)

If tstat < cv, we decide to reject H_0

-1.53 > -2.86, Ho cannot be rejected.
Thus: There is a unit root

- Conclusion 1: no lags on Δy is needed.
- Conclusion 2: the *t*-statistic of the parameter on y_{t-1} is -1.53, which is above the critical value of –2.86 in the DF-table.
  Conclusion: H_0 cannot be rejected. *AEXindex* has a unit root.

2. Test for *stockprice* (time trend)

```
. reg d.AEXindex  l.d.AEXindex l2.d.AEXindex l3.d.AEXindex l.AEXindex day
      Source |       SS       df       MS              Number of obs =     506
-------------+------------------------------           F(  5,   500) =    1.23
       Model |  1304.99105     5   260.99821           Prob > F      =  0.2930
    Residual |  105966.262   500  211.932524           R-squared     =  0.0122
-------------+------------------------------           Adj R-squared =  0.0023
       Total |  107271.253   505  212.418323           Root MSE      =  14.558
------------------------------------------------------------------------------
  D.AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    AEXindex |
         LD. |   .0459267   .0452216     1.02   0.310     -.042921    .1347745
        L2D. |  -.0264859   .0453212    -0.58   0.559    -.1155293    .0625576
        L3D. |   .0508838   .0454005     1.12   0.263    -.0383155    .1400831
         L1. |  -.0178597   .0091222    -1.96   0.051    -.0357822    .0000629
         day |    .009889   .0077735     1.27   0.204    -.0053839    .0251618
       _cons |   20.56107   10.07844     2.04   0.042     .7597643    40.36237
------------------------------------------------------------------------------

. test  l.d.AEXindex l2.d.AEXindex l3.d.AEXindex

 ( 1)  LD.AEXindex = 0
 ( 2)  L2D.AEXindex = 0
 ( 3)  L3D.AEXindex = 0
       F(  3,   500) =    0.83
            Prob > F =    0.4782
```

Estimated parameter on day: The current period´s change in the AEXindex increases by 0.0099 per day, holding 3 periods lagged differences constant.

```
. reg d.AEXindex  l.d.AEXindex l.AEXindex day
      Source |       SS       df       MS              Number of obs =     508
-------------+------------------------------           F(  3,   504) =    1.53
       Model |  966.988196     3  322.329399           Prob > F      =  0.2065
    Residual |  106377.635   504  211.066737           R-squared     =  0.0090
-------------+------------------------------           Adj R-squared =  0.0031
       Total |  107344.624   507  211.725096           Root MSE      =  14.528
------------------------------------------------------------------------------
  D.AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    AEXindex |
         LD. |   .0425579   .0450624     0.94   0.345    -.0459753    .1310911
         L1. |  -.0173047   .0089625    -1.93   0.054    -.0349132    .0003037
         day |   .0094712   .0076246     1.24   0.215    -.0055088    .0244511
       _cons |   19.96979   9.918958     2.01   0.045     .4821953    39.45739
------------------------------------------------------------------------------
```

```
. reg d.AEXindex  l.AEXindex day
      Source |       SS       df       MS              Number of obs =     509
-------------+------------------------------           F(  2,   506) =    1.85
       Model |  778.508723     2  389.254362           Prob > F      =  0.1587
    Residual |  106600.927   506  210.673769           R-squared     =  0.0073
-------------+------------------------------           Adj R-squared =  0.0033
       Total |  107379.436   508  211.376842           Root MSE      =  14.515
------------------------------------------------------------------------------
  D.AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    AEXindex |
         L1. |  -.0162843   .0088746    -1.83   0.067    -.0337199    .0011513
         day |   .0087559   .0075405     1.16   0.246    -.0060587    .0235704
       _cons |   18.85052   9.831666     1.92   0.056    -.4653963    38.16643
------------------------------------------------------------------------------
```

Conclusion of unit root does not change.

3. Test for *stockprice* (no time trend)

```
. reg d.stockprice l.d.stockprice l2.d.stockprice l3.d.stockprice l.stockprice

      Source |       SS       df       MS              Number of obs =     506
-------------+------------------------------           F(  4,   501) =    0.47
       Model |  21268.5588     4  5317.13969           Prob > F      =  0.7589
    Residual |  5686521.35   501   11350.342           R-squared     =  0.0037
-------------+------------------------------           Adj R-squared = -0.0042
       Total |  5707789.91   505  11302.5543           Root MSE      =  106.54
------------------------------------------------------------------------------
D.stockprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |
         LD. |  -.0354153   .0447523    -0.79   0.429    -.1233407      .05251
        L2D. |  -.0016895   .0449451    -0.04   0.970    -.0899935    .0866146
        L3D. |    .034884   .0450086     0.78   0.439     -.053545    .1233129
         L1. |  -.0028721   .0035604    -0.81   0.420    -.0098673    .0041231
             |
       _cons |   28.99591   26.70869     1.09   0.278    -23.47893    81.47075
------------------------------------------------------------------------------

. test l.d.stockprice l2.d.stockprice l3.d.stockprice

 ( 1)  LD.stockprice = 0
 ( 2)  L2D.stockprice = 0
 ( 3)  L3D.stockprice = 0

       F(  3,   501) =    0.41
            Prob > F =    0.7462
```

```
. reg d.stockprice l.d.stockprice  l.stockprice
      Source |       SS       df       MS              Number of obs =     508
-------------+------------------------------           F(  2,   505) =    0.60
       Model |  13523.8227     2  6761.91134           Prob > F      =  0.5496
    Residual |  5698814.53   505  11284.7813           R-squared     =  0.0024
-------------+------------------------------           Adj R-squared = -0.0016
       Total |  5712338.36   507  11266.9396           Root MSE      =  106.23
------------------------------------------------------------------------------
D.stockprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |
         LD. |  -.0350543   .0445508    -0.79   0.432     -.122582    .0524734
         L1. |  -.0025621   .0035289    -0.73   0.468    -.0094953     .004371
       _cons |   26.78994   26.47965     1.01   0.312     -25.2339    78.81378
------------------------------------------------------------------------------

. reg d.stockprice l.stockprice
      Source |       SS       df       MS              Number of obs =     509
-------------+------------------------------           F(  1,   507) =    0.49
       Model |  5486.72812     1  5486.72812           Prob > F      =  0.4859
    Residual |  5720620.49   507  11283.2751           R-squared     =  0.0010
-------------+------------------------------           Adj R-squared = -0.0010
       Total |  5726107.22   508  11271.8646           Root MSE      =  106.22
------------------------------------------------------------------------------
D.stockprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |
         L1. |  -.0024544   .0035196    -0.70   0.486    -.0093692    .0044605
       _cons |   25.50761   26.41139     0.97   0.335    -26.38164    77.39687
------------------------------------------------------------------------------
```

- Conclusion 1: no lags on Δy is needed.
- Conclusion 2: the *t*-statistic of the parameter on y_{t-1} is -0.70, which is above the critical value of –2.86 in the DF-table.
  Conclusion: H_0 cannot be rejected. *Stockprice* has a unit root.

4. **Test for *stockprice* (time trend)**

```
. reg d.stockprice l.d.stockprice l2.d.stockprice l3.d.stockprice l.stockprice
day

      Source |       SS       df       MS              Number of obs =     506
-------------+------------------------------           F(  5,   500) =    2.80
       Model |  155454.684     5  31090.9367           Prob > F      =  0.0166
    Residual |  5552335.23   500  11104.6705           R-squared     =  0.0272
-------------+------------------------------           Adj R-squared =  0.0175
       Total |  5707789.91   505  11302.5543           Root MSE      =  105.38

------------------------------------------------------------------------------
D.stockprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |
         LD. | -.0116868   .0447886    -0.26   0.794    -.0996838    .0763101
        L2D. |  .0207671   .0449229     0.46   0.644    -.0674939    .1090281
        L3D. |  .0555125   .0449127     1.24   0.217    -.0327283    .1437533
         L1. | -.0494445   .0138527    -3.57   0.000    -.0766613   -.0222277
             |
         day |  .4386184   .1261787     3.48   0.001     .1907127    .6865241
       _cons |  259.9198   71.49079     3.64   0.000     119.4604    400.3792
------------------------------------------------------------------------------

. test l.d.stockprice l2.d.stockprice l3.d.stockprice

 ( 1)  LD.stockprice = 0
 ( 2)  L2D.stockprice = 0
 ( 3)  L3D.stockprice = 0

       F(  3,   500) =    0.59
            Prob > F =    0.6206
```

```
. reg d.stockprice l.d.stockprice l.stockprice day

      Source |       SS       df       MS              Number of obs =     508
-------------+------------------------------           F(  3,   504) =    4.19
       Model |  138911.355     3  46303.7848           Prob > F      =  0.0061
    Residual |     5573427   504  11058.3869           R-squared     =  0.0243
-------------+------------------------------           Adj R-squared =  0.0185
       Total |  5712338.36   507  11266.9396           Root MSE      =  105.16

------------------------------------------------------------------------------
D.stockprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |
         LD. |  -.0138589   .0445486    -0.31   0.756    -.1013827    .0736648
         L1. |  -.0464675   .0134986    -3.44   0.001     -.072988    -.019947
             |
         day |   .4139798   .1229413     3.37   0.001     .1724392    .6555203
       _cons |   244.7914   69.84611     3.50   0.000      107.566    382.0168
------------------------------------------------------------------------------
```

```
. reg d.stockprice l.stockprice day

      Source |       SS       df       MS              Number of obs =     509
-------------+------------------------------           F(  2,   506) =    6.50
       Model |  143367.438     2  71683.7191           Prob > F      =  0.0016
    Residual |  5582739.78   506  11033.0826           R-squared     =  0.0250
-------------+------------------------------           Adj R-squared =  0.0212
       Total |  5726107.22   508  11271.8646           Root MSE      =  105.04

------------------------------------------------------------------------------
D.stockprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |
         L1. |  -.0478543   .0133058    -3.60   0.000    -.0739958   -.0217129
             |
         day |   .4282322   .1211368     3.54   0.000     .1902392    .6662252
       _cons |   251.1049   68.95356     3.64   0.000     115.6343    386.5754
------------------------------------------------------------------------------
```

- Conclusion 1: no lags on Δy is needed.
- Conclusion 2: the *t*-statistic of the parameter on y_{t-1} is -3.60, which is below the critical value of –3.41 in the DF-table, but it is above the critical value of -3.96.
  Conclusion: H_0 can be rejected. *Stockprice* has no unit root for significance level of 0.05 and *Stockprice* has a unit root for significance level of 0.01.
  In what follows, we assume that Stockprice has a unit root.

Co-integration

Co-integration and error-correction models

Aim: to introduce a co-integrating regression

- Co-integration theory was developed by Engle and Granger (1987)
- Starting point: two variables that are both I(1): y_t and x_t.
  - A Dickey Fuller test indicates whether the variables are I(0) or I(1).
- One would expect that linear combinations of y_t and x_t are non-stationary also.
- Both variables are co-integrated if there is a β, for which the linear combination of y_t and x_t is stationary (I(0)). The linear combination is $y_t - \beta x_t$
- Economic interpretation of co-integration: the difference between y_t and βx_t is on average constant. There is a tendency to move back to this constant. Also, the variance of $y_t - \beta x_t$ is constant. Thus, $y_t = \alpha + \beta x_t$ is a long-run relationship.
- There are two cases where co-integration occurs.
- Case 1: β is known. Compute $s_t = y_t - \beta x_t$ and test whether there is a unit root in s_t using a Dickey Fuller test.
  - If the null hypothesis is rejected (of a unit root in s_t) then y_t and x_t will be co-integrated.
- Case 2: β is unknown and must be estimated. If y_t and x_t are co-integrated the OLS estimator $\hat{\beta}$ from the co-integrating regression:

$$\hat{y}_t = \hat{\alpha} + \hat{\beta} x_t \qquad (18.31)$$

is consistent for β.

  - Determine the residual: $\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta} x_t$

- Since $\hat{\beta}$ is estimated, the critical values of Tables 18.2 and 18.3 cannot be used.

Testing for co-integration (after co-integrating regression (18.31) (Engle-Granger test):

Table 18.4

Signif. level	1%	2.5%	5%	10%
Critical value	-3.90	-3.59	-3.34	-3.04

- The model: $y_t = \alpha + \beta x_t$
- Dickey Fuller test on co-integration of y_t and x_t. We run a regression of $\Delta \hat{u}_t$ on $\hat{u}_{t-1}$, and we may compare the *t*-statistic on $\hat{u}_{t-1}$ with the critical value in table 18.3.
    - If fail to reject H_0: regression of y_t on x_t is spurious.
    - If reject H_0: regression of y_t on x_t is a co-integrating regression.
- There may be a time-trend in the co-integrating regression:

$$y_t = \hat{\alpha} + \hat{\eta} t + \hat{\beta} x_t \tag{18.32}$$

Table 18.5

Signif. level	1%	2.5%	5%	10%
Critical value	-4.32	-4.03	-3.78	-3.50

Example 2: test on co-integration

```
. reg AEXindex stockprice

      Source |       SS       df       MS              Number of obs =     510
-------------+------------------------------           F(  1,   508) = 1610.23
       Model |  6031609.59     1  6031609.59           Prob > F      =  0.0000
    Residual |  1902875.37   508  3745.81766           R-squared     =  0.7602
-------------+------------------------------           Adj R-squared =  0.7597
       Total |  7934484.96   509  15588.3791           Root MSE      =  61.203

------------------------------------------------------------------------------
    AEXindex |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  stockprice |   .0811894   .0020233    40.13   0.000     .0772144    .0851645
       _cons |   675.5636   15.19135    44.47   0.000      645.718    705.4092
------------------------------------------------------------------------------

. predict uhat, resid

. reg d.uhat l.uhat

      Source |       SS       df       MS              Number of obs =     509
-------------+------------------------------           F(  1,   507) =    1.74
       Model |  195.994111     1  195.994111           Prob > F      =  0.1881
    Residual |  57211.1212   507  112.842448           R-squared     =  0.0034
-------------+------------------------------           Adj R-squared =  0.0014
       Total |  57407.1153   508  113.006133           Root MSE      =  10.623

------------------------------------------------------------------------------
      D.uhat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        uhat |
         L1. |  -.0101998   .0077394    -1.32   0.188    -.0254051    .0050054
             |
       _cons |  -.2724943    .470849    -0.58   0.563     -1.19755    .6525611
------------------------------------------------------------------------------
```

- The above output indicates that *AEXindex* and *stockprice* are not co-integrated, as the *t*-statistic on $\hat{u}_{t-1}$ (-1.32) is above the critical value (Table 18.4;   -3.34 at the 5%-level).

```
. reg d.uhat l.uhat day

      Source |       SS       df       MS              Number of obs =     509
-------------+------------------------------           F(  2,   506) =    1.42
       Model |  320.225932     2  160.112966           Prob > F      =  0.2429
    Residual |  57086.8894   506   112.81994           R-squared     =  0.0056
-------------+------------------------------           Adj R-squared =  0.0016
       Total |  57407.1153   508  113.006133           Root MSE      =  10.622

------------------------------------------------------------------------------
      D.uhat |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        uhat |
         L1. |  -.0106252   .0077493    -1.37   0.171    -.0258499    .0045995
             |
         day |  -.0033669   .0032085    -1.05   0.295    -.0096705    .0029368
       _cons |   .5895392   .9468343     0.62   0.534    -1.270671     2.44975
------------------------------------------------------------------------------
```

- The above output indicates that *AEXindex* and *stockprice* are not co-integrated, as the *t*-statistic on $\hat{u}_{t-1}$ (-1.37) is above the critical value (Table 18.5; -3.78 at the 5%-level).

Error-correction models

Error-correction models

Aim: to consider the economic implications of error-correction models.

Case 1 (no error correction): y_t and x_t are both I(1), and there is no co-integration between y_t and x_t (no rejection of H_0 after DF-test on residuals of (18.31). The model becomes:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + u_t \tag{18.36}$$

All variables are in first differences.

Case 2 (error-correction model): y_t and x_t are both I(1), and there is co-integration between y_t and x_t (rejection of H_0 after DF-test on residuals of (18.31)).

$$\begin{aligned}\Delta y_t &= \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \delta s_{t-1} + u_t = \\ &= \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \delta (y_{t-1} - \beta x_{t-1}) + u_t \end{aligned} \tag{18.37}$$

- $s_{t-1} = y_{t-1} - \beta x_{t-1}$ is called the error-correction term. Note that the lag *t*-1 of *s* is included in the equation.

Economic interpretation:

$$\Delta y_t = \alpha_0 + \gamma_0 \Delta x_t + \delta (y_{t-1} - \beta x_{t-1}) + u_t \tag{18.38}$$

where $\delta < 0$.

- Δy_t will return to equilibrium if $y_{t-1} > \beta x_{t-1}$ or $y_{t-1} < \beta x_{t-1}$ as $\delta < 0$. If $y_{t-1} > \beta x_{t-1}$ the return to equilibrium will involve a decrease in Δy_t, and if $y_{t-1} < \beta x_{t-1}$ an it will lead to increase in Δy_t. The adjustments are because the parameter $\delta < 0$
- In equation (18.38) a lag of Δy_t was excluded to keep it simple.

How is an error-correction model estimated?
Aim: to demonstrate the estimation procedure.

Case 1: β is known.
Regress Δy_t on Δx_t and s_{t-1}, where $s_{t-1} = y_{t-1} - \beta x_{t-1}$
s_{t-1} is the error correction term.
Note that the lag *t*-1 of *s* is included in the equation.

Case 2: β is unknown - two-stage procedure
Regress Δy_t on Δx_t and $\hat{s}_{t-1}$, where $\hat{s}_{t-1} = y_{t-1} - \hat{\beta} x_{t-1}$

Summary

- The model: $y_t = \alpha + \beta x_t$

  - Step 1: Dickey Fuller tests for unit root in y_t and x_t
    - This determines whether to include a time trend and how many lags of Δy_t are appropriate.
  - If both y_t and x_t are I(1) continue with the following:
  - Step 2: co-integrating regression
  - Step 3: Dickey-Fuller test (for unit root), to the residuals $\hat{u}_t = y_t - \hat{\alpha} - \hat{\beta} x_t$
  - Step 4a: If the model contains unit root (H0 of step 3 cannot be rejected), estimate the model in first differences.
  - Step 4b: If the model contains no unit root (H0 of step 3 can be rejected) use the error-correction model.
  - Step 5 (after 4b). Estimate the error-correction model using the two-stage procedure.

Forecasting

Section 18.5: Forecasting

Aim: to introduce forecasting procedures

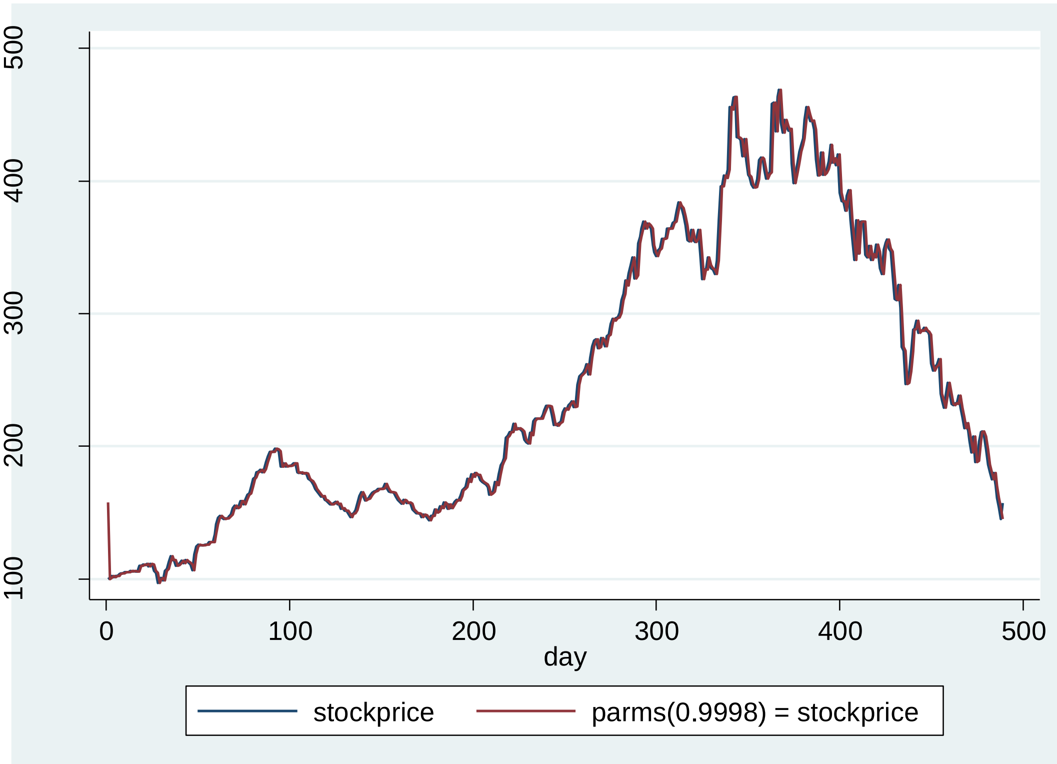
- Next we consider forecasting with time series, and we apply it to dynamic models.
- At time *t* we would like to make a forecast of *y* at *t*+1.
- Notation:
    - I_t: all information available at time *t*.
    - f_t is the one-period ahead forecast.
    - $e_{t+1} = y_{t+1} - f_t$: forecast error
- We consider the squared error (as a loss function)
    - $E(e_{t+1}^2 \mid I_t) = E[(y_{t+1} - f_t)^2 \mid I_t]$ (18.40)
    - The squared error is minimal at $f_t = E(y_{t+1} \mid I_t)$
    - Best forecast: $E(y_{t+1} \mid I_t)$
- Example 3: Forecast based on a martingale (it means that adding more lagged variables does not provide more information to the equation):
    - $E(y_{t+1} \mid y_t, y_{t-1}, ..., y_0) = y_t$
- Example 4: Forecast based on **exponential smoothing**:
    - $f_t = \alpha y_t + (1-\alpha) f_{t-1}$ (with $f_0 = y_0$ and $0 < \alpha < 1$) so that the forecast of y_{t+1} is a weighted average of the previous forecast (at *t*-1) and the current realization of *y*).
    - $E(y_{t+1} \mid I_t) = \alpha y_t + \alpha(1-\alpha) y_{t-1} + ... + \alpha(1-\alpha)^t y_0$
- The application of smoothing (in addition to forecasting) can be used to extract an overall trend from a time-series. This is because a time series may contain a trend component as well as an irregular component. The researcher might only be interested in the trend component.

Example 5: stockdata.dta

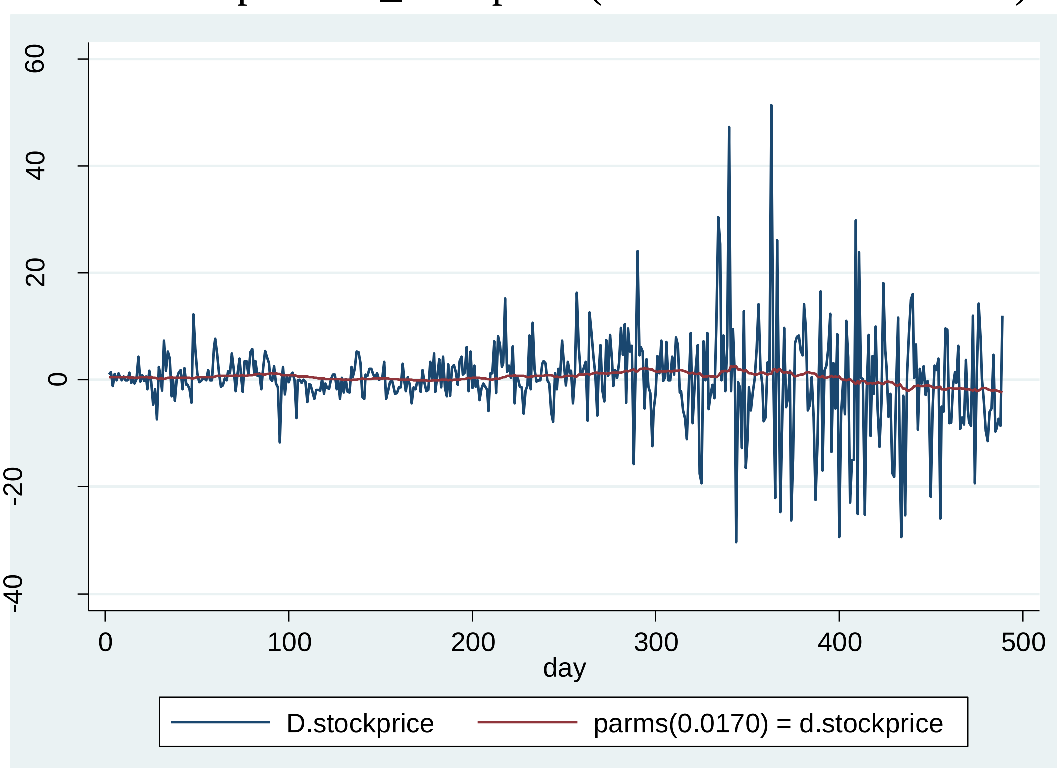
- Example 1 showed that *stockprice* is I(1). It contains a unit root, so the level of *stockprice* is uninformative.

```
. tssmooth exponential lev_stockprice = stockprice

computing optimal exponential  coefficient (0,1)

optimal exponential coefficient =        0.9998
sum-of-squared residuals        =     36603.576
root mean squared error         =      8.651817

. tssmooth exponential dif_stockprice = d.stockprice

computing optimal exponential  coefficient (0,1)

optimal exponential coefficient =        0.0170
sum-of-squared residuals        =     33045.423
root mean squared error         =     8.2289751
. tsline gfr gfr_smooth_level
. tsline d.gfr gfr_smooth_dif

(see graphs below)
```

tsline stockprice lev_stockprice



. tsline d.stockprice dif_stockprice (on first-differenced data)



Using regression models for forecasting

Aim: to show that regressions can be used for forecasts and forecast intervals.

- Below is a dynamic regression equation, in which the forecast depends on lagged values of *y* and *z*:
  $y_t = \delta_0 + \alpha_1 y_{t-1} + \gamma_1 z_{t-1} + u_t$
  $E(u_t \mid I_{t-1}) = 0$
- The forecast of y_{t+1} is $\delta_0 + \alpha_1 y_t + \gamma_1 z_t$
- Thus: $\hat{f}_t = \hat{\delta}_0 + \hat{\alpha}_1 y_t + \hat{\gamma}_1 z_t$
- The forecast error is $e_{t+1} = y_{t+1} - \hat{f}_t$
- The variance of the forecast error is:
  $Var(e_{t+1}) = Var(u_t) + Var(\hat{f}_t) = \hat{\sigma}_u^2 + Var(\hat{f}_t)$
- A 95% forecast interval: $\hat{f}_t \pm 2 * \sqrt{\hat{\sigma}_u^2 + Var(\hat{f}_t)}$ (18.47)

Vector autoregressive (VAR) model

Vector autoregressive (VAR) model: Sims (Nobel prize, 2011)

Aim: to introduce VAR-models for forecasting.

- Model:

$$y_t = \delta_0 + \alpha_1 y_{t-1} + \gamma_1 z_{t-1} + \alpha_2 y_{t-2} + \gamma_2 z_{t-2} + \ldots \tag{18.50}$$
$$z_t = \eta_0 + \beta_1 y_{t-1} + \rho_1 z_{t-1} + \beta_2 y_{t-2} + \rho_2 z_{t-2} + \ldots$$

- A VAR model may be applied for forecasting

e.g.

$$y_t = \delta_0 + \alpha_1 y_{t-1} + \gamma_1 z_{t-1} + u_{1t}$$
$$z_t = \eta_0 + \beta_1 y_{t-1} + \rho_1 z_{t-1} + u_{2t}$$

u_{1t} and u_{2t} are i.i.d. (with different variances and they are mutually independent). They are independent of all previous observations of *y* and *z*.

- Three advantages of VAR models are that:
  a) The components can be considered simultaneously
  b) The model can be more parsimonious
  c) The model may include fewer lags

This allows more accurate forecasting, because the information set (the right-hand side variables of the regression equation) is extended and also includes the history of other variables.

- Impulse Response Functions (IRF) is also used to produce forecasts. (no part of this course)
- The error terms are assumed to be stationary (no unit root).
- (18.50) is not a structural equation model because:
  - The exclusion constraints (see forthcoming lecture 7) are not necessary to identify the structural model.
  - There is no a priori distinction between endogenous and exogenous variables.
- OLS on equations of equation (18.50) separately will yield consistent estimates.
- It is not easy to select the optimal lag length. Estimate the model and then apply the Akaike Information Criterion (AIC, not part of this course). Choose the model with the lowest AIC. The AIC is a selection criterion that is comparable to the adjusted *R*-

squared (remember that the selected model is often the one with the highest *R*-squared).

Example 6: VAR

The command for VAR (we apply 2 lags)

```
. var d.AEXindex d.stockprice, lags(1/2)

Vector autoregression

Sample:  4 - 510                                    No. of obs      =       507
Log likelihood = -4994.471                          AIC             =   19.7415
FPE            =   1284333                          HQIC            =  19.77421
Det(Sigma_ml)  =   1234654                          SBIC            =  19.82491

Equation           Parms      RMSE     R-sq      chi2     P>chi2
----------------------------------------------------------------
D_AEXindex           5     14.5954   0.0035   1.766256   0.7786
D_stockprice         5     106.408   0.0050   2.529515   0.6394
----------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
D_AEXindex   |
    AEXindex |
         LD. |   .0145237   .0617081     0.24   0.814     -.106422    .1354694
        L2D. |  -.0643886   .0615668    -1.05   0.296    -.1850573      .05628
             |
  stockprice |
         LD. |   .0037053   .0084287     0.44   0.660    -.0128147    .0202353
        L2D. |   .0059473   .0084252     0.71   0.480    -.0105657    .0224604
             |
       _cons |   .2688697   .6495078     0.41   0.679    -1.004142    1.541882
-------------+----------------------------------------------------------------
D_stockprice |
    AEXindex |
         LD. |   .5195352   .4498833     1.15   0.248      -.36222     1.40129
        L2D. |  -.3051594   .4488529    -0.68   0.497    -1.184895    .5745761
             |
  stockprice |
         LD. |  -.0839891   .0614497    -1.37   0.172    -.2044284    .0364502
        L2D. |   .0207528   .0614238     0.34   0.735    -.0996357    .1411413
             |
       _cons |   8.039581   4.735239     1.70   0.090    -1.241318    17.32048
------------------------------------------------------------------------------
```

Wrapping up

We have considered:

- Testing for unit roots
- Co-integration and error correction models
- A test for co-integration: Engle and Granger test.
- Error-correction
- Prediction and smoothing
- Vector Autoregressive (VAR) models