Autoregressive Integrated Moving Average (ARIMA) Models

Example

MA(1)

$$Y_{1t} = 0.1 + u_{1t} + 0.5\, u_{1t-1}$$

AR(1)

$$Y_{2t} = 0.1 + 0.5\, Y_{2t-1} + u_{2t}$$

ARMA(1,1)

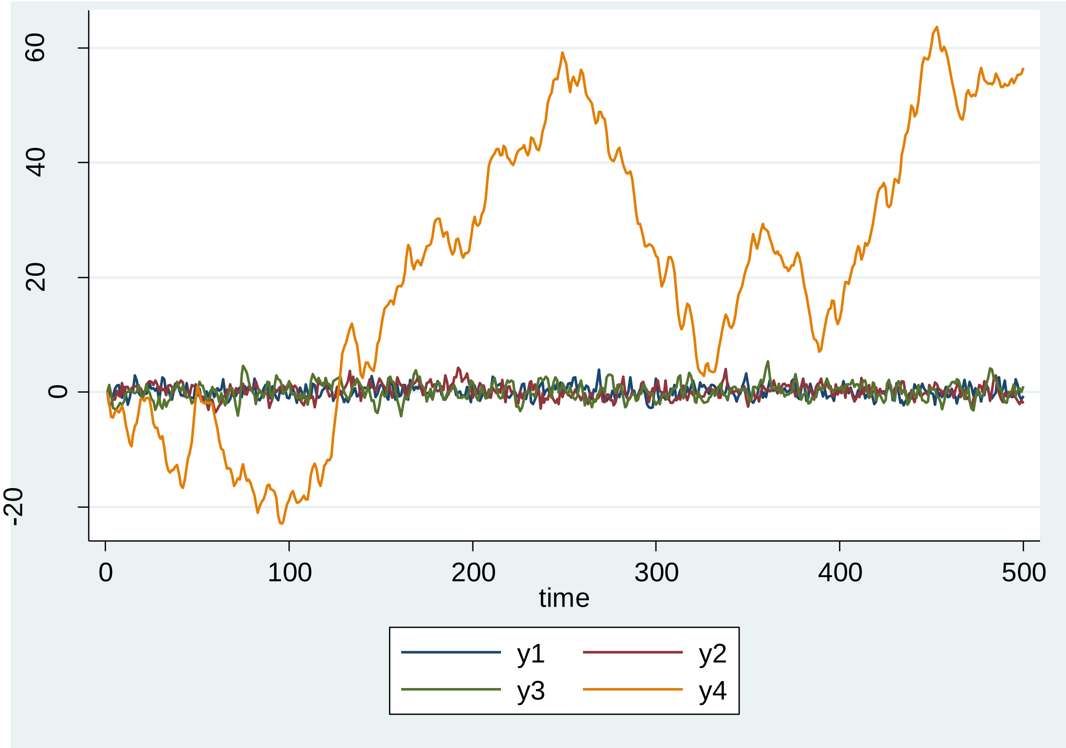
$$Y_{3t} = 0.1 + 0.5\, Y_{3t-1} + u_{3t} + 0.5\, u_{3t-1}$$

ARIMA(1,1,1)

$$\Delta Y_{4t} = 0.1 + 0.5\, \Delta Y_{4t-1} + u_{4t} + 0.5\, u_{4t-1}$$

```
. set obs 500
obs was 0, now 500

. g time=_n

. tsset time
        time variable:  time, 1 to 500
                delta:  1 unit

. set seed 5432

. g y1=0 in 1
(499 missing values generated)

. g y2=0 in 1
(499 missing values generated)

. g y3=0 in 1
(499 missing values generated)

. g y4=0 in 1
(499 missing values generated)

. g dy4 = 0 in 1
(499 missing values generated)

. g u1= rnormal(0,1)

. g u2= rnormal(0,1)

. g u3= rnormal(0,1)

. g u4= rnormal(0,1)

. replace y1=0.1+u1+0.5*l.u1 if time>1
(499 real changes made)

. replace y2=0.1+0.5*l.y2+u2 if time>1
(499 real changes made)

. replace y3=0.1+0.5*l.y3+u3+0.5*l.u3 if time>1
(499 real changes made)

. replace y4=0.1+l.y4+u4+0.5*l.u4 if time>1
(499 real changes made)

. replace dy4=d.y4 if time>1
(499 real changes made)

. replace y4=0.1+l.y4+0.5*l.dy4+u4+0.5*l.u4 if time>1
(498 real changes made)
```

```
. line y1 y2 y3 y4 time
```



Example: Finding ARIMA process – Box-Jenkins Methodology

Identification and Estimation: The first step is to identify the order of ARIMA(p,d,q). To find the value of p, d, and q, we should firstly identify value of d by performing unit root test to find which order the series is integrated.

Unit Root Test

Find integrated order of the series using unit root test.

```
. tsset time
        time variable:  time, 1 to 357
                delta:  1 unit

. dfuller y, trend lags(1) regress

Augmented Dickey-Fuller test for unit root         Number of obs   =       355

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.152            -3.986            -3.426            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.5167

------------------------------------------------------------------------------
D.y          |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |  -.0237778   .0110475    -2.15   0.032    -.0455054   -.0020502
         LD. |   .0625276   .0531988     1.18   0.241    -.0421009     .167156
      _trend |   .0143903   .0082343     1.75   0.081    -.0018044     .030585
       _cons |   23.00759   11.00167     2.09   0.037     1.370103    44.64508
```

```
------------------------------------------------------------------------------

. dfuller y,  lags(1) regress

Augmented Dickey-Fuller test for unit root         Number of obs   =       355

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.414            -3.452            -2.876            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.5757

------------------------------------------------------------------------------
         D.y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |  -.0129893   .0091884    -1.41   0.158    -.0310604    .0050818
         LD. |    .059847   .0533316     1.12   0.263    -.0450416    .1647357
             |
       _cons |   14.12272   9.784939     1.44   0.150    -5.121575    33.36702
------------------------------------------------------------------------------

. dfuller d.y, trend lags(1) regress

Augmented Dickey-Fuller test for unit root         Number of obs   =       354

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)            -12.054            -3.986            -3.426            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000

------------------------------------------------------------------------------
D2.y         |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         D.y |
         L1. |  -.8856237   .0734696   -12.05   0.000    -1.030121   -.7411262
         LD. |  -.0668939   .0533317    -1.25   0.211    -.1717849    .0379971
      _trend |   .0042522   .0069007     0.62   0.538    -.0093198    .0178243
       _cons |  -.4529247   1.418181    -0.32   0.750    -3.242154    2.336305
------------------------------------------------------------------------------

. arima y, arima(1,1,1) nolog

ARIMA regression

Sample:  2 - 357                                Number of obs      =       356
                                                Wald chi2(2)       =      3.13
Log likelihood = -1423.146                      Prob > chi2        =    0.2089

------------------------------------------------------------------------------
             |                 OPG
         D.y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
       _cons |   .3700177   .8218557     0.45   0.653     -1.24079    1.980825
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .2948697   .6818072     0.43   0.665    -1.041448    1.631187
             |
          ma |
         L1. |  -.2368839   .7055342    -0.34   0.737    -1.619705    1.145938
-------------+----------------------------------------------------------------
      /sigma |   13.17852   .3681099    35.80   0.000     12.45704        13.9
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
confidence interval is truncated at zero.

. est store arima111

. qui arima y, arima(1,1,2) nolog
```

```
. est store arima112

. qui arima y, arima(1,1,3) nolog

. est store arima113

. qui arima y, arima(2,1,1) nolog

. est store arima211

. qui arima y, arima(2,1,2) nolog

. est store arima212

. qui arima y, arima(2,1,3) nolog

. est store arima213

. qui arima y, arima(3,1,1) nolog

. est store arima311

. qui arima y, arima(3,1,2) nolog

. est store arima312

. qui arima y, arima(3,1,3) nolog

. est store arima313

. est table arima11*, star(0.1 0.05 0.01) stat(N ll chi2 aic bic)

--------------------------------------------------------------
    Variable |   arima111        arima112        arima113
-------------+------------------------------------------------
y            |
       _cons |  .37001774       .37418642        .3987889
-------------+------------------------------------------------
ARMA         |
          ar |
         L1. |  .29486971      -.31289111       .97476406***
             |
          ma |
         L1. | -.23688394       .36887941      -.92682332
         L2. |                  .11337558**     .03029933
         L3. |                                 -.10347864
-------------+------------------------------------------------
sigma        |
       _cons |  13.178521***    13.127234***    13.055651
-------------+------------------------------------------------
Statistics   |
           N |        356             356             356
          ll | -1423.1463      -1421.7567      -1420.6373
        chi2 |  3.1319865       6.7574223       9222.4937
         aic |  2854.2926       2853.5135       2853.2745
         bic |  2869.7923       2872.8881       2876.5241
--------------------------------------------------------------
                        legend: * p<.1; ** p<.05; *** p<.01

. est table arima21*, star(0.1 0.05 0.01) stat(N ll chi2 aic bic)

--------------------------------------------------------------
    Variable |   arima211        arima212        arima213
-------------+------------------------------------------------
y            |
       _cons |  .37274811       .37283121       .37413313
-------------+------------------------------------------------
ARMA         |
          ar |
         L1. | -.28216023      -.66642562**     .80124027**
         L2. |  .09539459**    -.39627716       -.3650645
             |
          ma |
         L1. |  .33242601       .71235711***   -.75250981**
         L2. |                  .50413388       .39755506
         L3. |                                 -.10013858*
-------------+------------------------------------------------
```

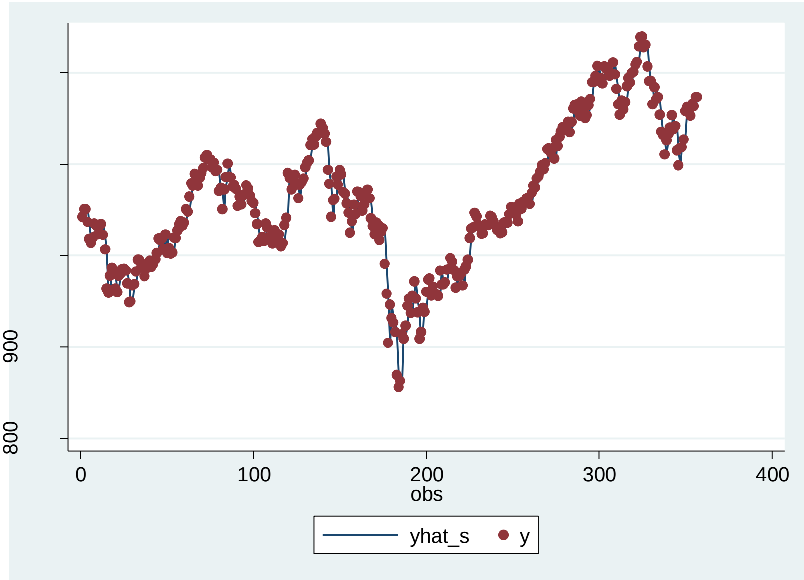
```
sigma        |
       _cons |  13.139922***    13.112555***    13.079159***
-------------+------------------------------------------------
Statistics   |
           N |        356             356             356
          ll | -1422.0873      -1421.3508       -1420.453
        chi2 |  4.7048993       17.316358       18.662816
         aic |  2854.1745       2854.7016       2854.9059
         bic |  2873.5492       2877.9512       2882.0304
--------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01

. est table arima31*, star(0.1 0.05 0.01) stat(N ll chi2 aic bic)

--------------------------------------------------------------
    Variable |   arima311        arima312        arima313
-------------+------------------------------------------------
y            |
       _cons |  .37269062        .3751536       .37488954
-------------+------------------------------------------------
ARMA         |
          ar |
         L1. |  .51072432       .67284808*      .63119128
         L2. |  .04992461      -.20028258      -.15281824
         L3. | -.12395064***   -.11233502**    -.16502621
             |
          ma |
         L1. | -.46077982      -.62568637*     -.58388774
         L2. |                   .2450268       .19971418
         L3. |                                  .05258667
-------------+------------------------------------------------
sigma        |
       _cons |  13.082025***    13.075786***    13.075313***
-------------+------------------------------------------------
Statistics   |
           N |        356             356             356
          ll | -1420.5393      -1420.3504      -1420.3353
        chi2 |  11.701303       16.675786       16.501777
         aic |  2853.0786       2854.7008       2856.6706
         bic |  2876.3282       2881.8253       2887.67
--------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01
```

Forecast

Command: predict, xb

```
. arima y, arima(1,1,1) nolog

ARIMA regression

Sample:  2 - 357                                Number of obs      =       356
                                                Wald chi2(2)       =      3.13
Log likelihood = -1423.146                      Prob > chi2        =    0.2089

------------------------------------------------------------------------------
             |                 OPG
         D.y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
       _cons |   .3700177   .8218557     0.45   0.653     -1.24079    1.980825
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .2948697   .6818072     0.43   0.665    -1.041448    1.631187
             |
          ma |
         L1. |  -.2368839   .7055342    -0.34   0.737    -1.619705    1.145938
-------------+----------------------------------------------------------------
      /sigma |   13.17852   .3681099    35.80   0.000     12.45704        13.9
------------------------------------------------------------------------------
. predict dyhat, xb

. g yhat_s=l.y+dyhat if time>1
(3 missing values generated)
```

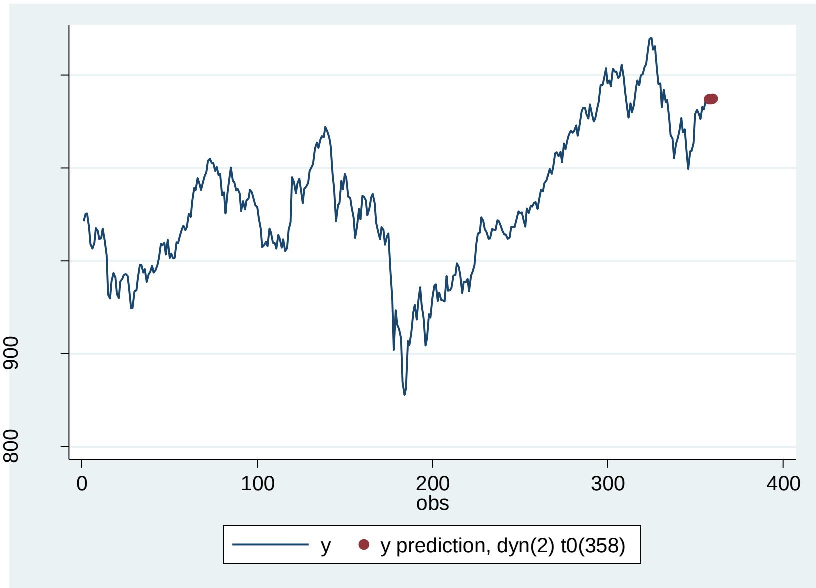
```
. twoway (line yhat_s time, sort) (scatter y time, sort)
```



Command: predict, y

Dynamic Forecast

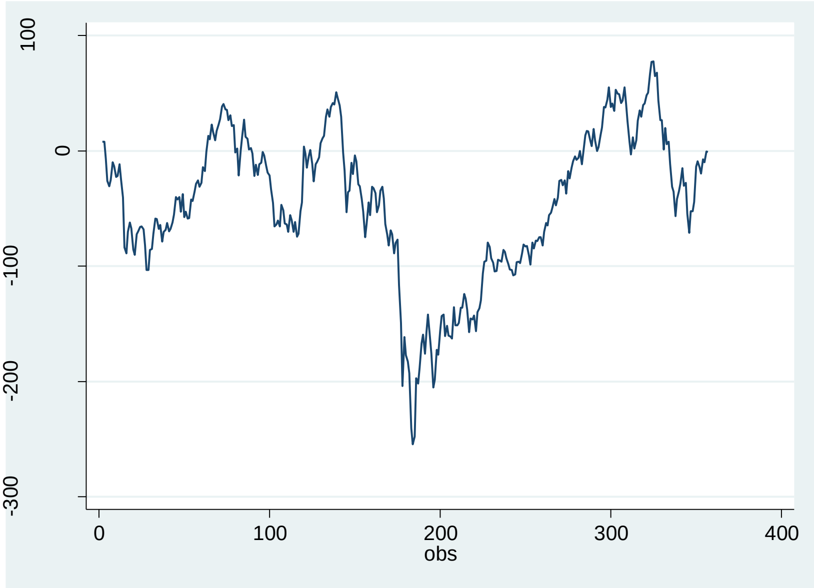
```
. predict yhat, y dynamic(.) t0(357)
Note: beginning dynamic predictions in period 2
(356 missing values generated)

. predict yhat2, y dynamic(.) t0(358)
Note: beginning dynamic predictions in period 2
(357 missing values generated)

. twoway (line y time, sort) (scatter yhat2 time, sort)
```



Forecasting Error

```
. predict yhat_d, y dynamic(.)
Note: beginning dynamic predictions in period 2
(1 missing value generated)

. g ferror=y-yhat_d if time<358
(4 missing values generated)
```

```
. twoway (line ferror time, sort)
```



SARIMA

```
. arima y, arima(1,1,1) sarima(1,1,1,13) nolog

ARIMA regression

Sample:  15 - 357                               Number of obs      =       343
                                                Wald chi2(4)       =    190.55
Log likelihood = -1390.897                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
      DS13.y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
       _cons |    .055653   .0949183     0.59   0.558    -.1303835    .2416896
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |  -.6610302   1.885136    -0.35   0.726    -4.355829    3.033769
             |
          ma |
         L1. |   .6497243   1.909893     0.34   0.734    -3.093598    4.393047
-------------+----------------------------------------------------------------
ARMA13       |
          ar |
         L1. |  -.0235579    .057453    -0.41   0.682    -.1361638     .089048
             |
          ma |
         L1. |  -.9555681   .0798977   -11.96   0.000    -1.112165   -.7989715
-------------+----------------------------------------------------------------
      /sigma |   13.33896   .5078398    26.27   0.000     12.34361    14.33431
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
confidence interval is truncated at zero.
```

ARMAX

```
. arima y d.usd, arima(1,1,1) nolog

ARIMA regression

Sample:   3 - 357                               Number of obs      =       355
                                                Wald chi2(3)       =     41.61
Log likelihood = -1405.373                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
```

```
         D.y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
         usd |
         D2. |  -25.89824   4.143719    -6.25   0.000    -34.01978    -17.7767
             |
       _cons |   .3479213   .8288965     0.42   0.675    -1.276686    1.972529
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .2162354   .3554955     0.61   0.543    -.4805229    .9129937
             |
          ma |
         L1. |  -.0959384   .3756771    -0.26   0.798    -.8322521    .6403752
-------------+----------------------------------------------------------------
      /sigma |   12.67782   .3774993    33.58   0.000     11.93794    13.41771
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
confidence interval is truncated at zero.
```