HG
Feb 2014 ECON 5101

Exercises III - 24 Feb 2014 (including answers)

Exercise 1.

In lecture notes 3 (LN3 – page 11) we estimated an ARMA(1,2) for $\Delta \ln BNP$ (Norwegian data) for the period, 1978q2 - 2013q2. Let $Y_t = \Delta \ln BNP_t$.

Model:

(1) $$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \text{ where } \varepsilon_t \sim WN(0, \sigma^2)$$

Table 1 Estimates (mle):

	φ_0	φ_1	θ_1	θ_2	σ
Estimate	0.0062	0.729	-1.105	0.490	0.0095
St.error	0.0012	0.124	0.116	0.084	0.00062

A. The mle estimates determine a causal stationary solution for Y_t. Why?

ANSWER: The root of the companion polynomial is $|1/0.729| > 1$.

B. The causal stationary solution is

(2) $$Y_t = \mu + \varepsilon_t + \delta_1 \varepsilon_{t-1} + \delta_2 \varepsilon_{t-2} + \cdots + \delta_j \varepsilon_{t-j} + \cdots$$

(i) Determine formulas for $\mu, \delta_1, \delta_2, \ldots, \delta_j, \ldots$

(ii) Determine the formula for the dynamic multiplier $\frac{\partial Y_{t+j}}{\partial \varepsilon_t}$

(iii) Determine the formula for the long run effect of a unit change in ε_t leaving the other errors terms unchanged. Estimate the long run effect from the data.

ANSWER:

(i) With lag polynomials (1) can be written

$$\varphi(L)Y_t = \varphi_0 + \theta(L)\varepsilon_t, \text{ where } \varphi(L) = 1 - \varphi_1 L \text{ and } \theta(L) = 1 + \theta_1 L + \theta_2 L^2$$

The causal stationary solution is

$$Y_t = \frac{\varphi_0}{\varphi(1)} + \frac{1+\theta_1 L + \theta_2 L^2}{1-\varphi_1 L}\varepsilon_t = \frac{\varphi_0}{1-\varphi_1} + \varepsilon_t + \delta_1\varepsilon_{t-1} + \delta_2\varepsilon_{t-2} + \cdots$$

so $\mu = \frac{\varphi_0}{1-\varphi_1}$. To find the $\delta's$ we can perform the multiplication (of companion power series)

$$\begin{aligned} 1+\theta_1 z+\theta_2 z^2 &= (1-\varphi_1 z)(1+\delta_1 z+\delta_2 z^2+\delta_3 z^3+\cdots) = \\ &= 1+\delta_1 z+\delta_2 z^2+\delta_3 z^3+\cdots \\ &\quad -\varphi_1 z-\varphi_1\delta_1 z^2-\varphi_1\delta_2 z^3-\cdots \\ &= 1+(\delta_1-\varphi_1)z+(\delta_2-\varphi_1\delta_1)z^2+(\delta_3-\varphi_1\delta_2)z^3+\cdots \end{aligned}$$

From the uniqueness of the power series expansion, we can identify coefficients on the left and right side, giving

$$\begin{aligned} &\delta_1-\varphi_1=\theta_1 \quad \Rightarrow \quad \delta_1=\theta_1+\varphi_1 \\ &\delta_2-\varphi_1\delta_1=\theta_2 \quad \Rightarrow \quad \delta_2=\theta_2+\varphi_1(\theta_1+\varphi_1) \\ &\delta_j-\varphi_1\delta_{j-1}=0 \text{ for } j=3,4,\ldots \quad \Rightarrow \quad \delta_j=\varphi_1\delta_{j-1}=\cdots=\varphi_1^{\,j-2}\delta_2 \text{ for } j=3,4,\ldots \end{aligned}$$

(ii) From the solution we get the dynamic multiplier, $\frac{\partial Y_{t+j}}{\partial\varepsilon_t}=\delta_j,\quad j=1,2,3,\ldots$

(iii) The long run effect is

$$\sum_{j=1}^{\infty}\frac{\partial Y_{t+j}}{\partial\varepsilon_t}=\sum_{j=1}^{\infty}\delta_j=\delta_1+\delta_2+\delta_2\varphi_1+\delta_2\varphi_1^2+\cdots=\delta_1+\delta_2\left(1+\varphi_1+\varphi_1^2+\cdots\right)=\delta_1+\frac{\delta_2}{1-\varphi_1}=$$

$$=\theta_1+\varphi_1+\frac{\theta_2+\varphi_1(\theta_1+\varphi_1)}{1-\varphi_1}=\frac{\theta_1+\theta_2+\varphi_1}{1-\varphi_1} \qquad \text{(Estimate: 0.421)}$$

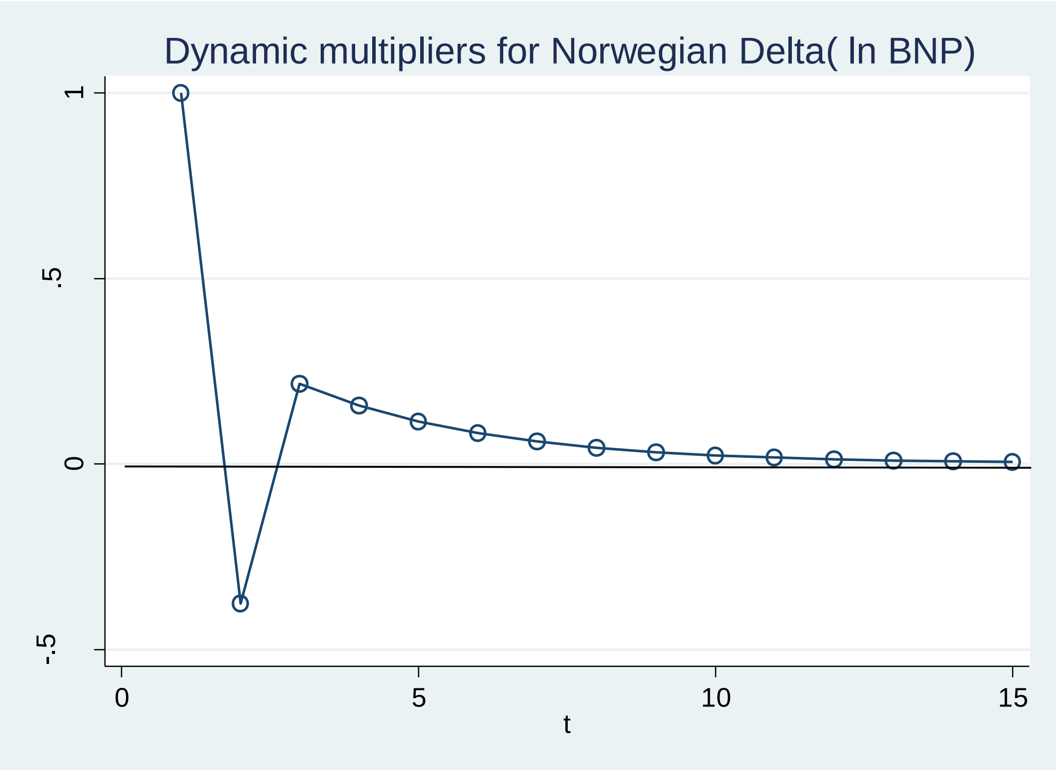
C.

Make a plot of the dynamic multipliers along the lines of Seminar I (appendix) – using the estimates in table 1. If you use Stata, use Stata 13.

ANSWER:

```
. set obs 30
obs was 0, now 30

. gen t=_n

. tsset t
```

```
        time variable:  t, 1 to 30
                delta:  1 unit

. gen delta=0

. replace delta = 1 in 1
(1 real change made)

. replace delta = -.376 in 2
(1 real change made)

. replace delta = .2159 in 3
(1 real change made)

. matrix delta=.729

. forecast create,replace
  Forecast model started.


. matrix coleq delta=delta:L.delta

. matrix list delta

symmetric delta[1,1]
     delta:
         L.
     delta
r1    .729

. forecast coefvector delta
  Forecast model now contains 1 endogenous variable.

. forecast solve, begin(4) log(off)

Computing dynamic forecasts for current model.
----------------------------------------------
Starting period:  4
Ending period:    30
Forecast prefix:  f_

Forecast 1 variable spanning 27 periods.
----------------------------------------

. list f_delta if t<=10

     +----------+
     |  f_delta |
     |----------|
  1. |        1 |
  2. |    -.376 |
  3. |    .2159 |
  4. | .1573911 |
  5. | .1147381 |
     |----------|
  6. | .0836441 |
  7. | .0609765 |
  8. | .0444519 |
  9. | .0324054 |
 10. | .0236236 |
     +----------+

. twoway (tsline f_delta if t<=15, recast(connected) msize(medlarge)
msymbol(circle_hollow))
```



D. (*Potential trap in Hamilton*). An alternative method of finding the dynamic multipliers $\{\delta_j\}$, is using Hamilton sec. 1.2 (see page 13), who finds explicit formulas for the dynamic multipliers $\frac{\partial Y_{t+j}}{\partial w_t}$.

In our case $w_t = \varphi_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$, so $\frac{\partial Y_{t+j}}{\partial w_t}$ is hardly the dynamic multiplier we really want (i.e., $\frac{\partial Y_{t+j}}{\partial \varepsilon_t}$).

(i) Explain why $\frac{\partial Y_{t+j}}{\partial w_t}$ and $\frac{\partial Y_{t+j}}{\partial \varepsilon_t}$ are *not* the same thing in our case.

(ii) How would you use Hamiltons's solution to derive the relevant d.m.'s $\left\{\frac{\partial Y_{t+j}}{\partial \varepsilon_t}\right\}$?

ANSWER:

(i) ε_t occurs in several of the w_t's (i.e., in w_{t+2}, w_{t+1}, and w_t):

(ii) Let $\delta_{w,j} = \frac{\partial Y_{t+j}}{\partial w_t}$ be the Hamilton multipliers. Then

$$Y_{t+j} = w_{t+j} + \delta_{w,1} w_{t+j-1} + \cdots + \delta_{w,j-2} w_{t+2} + \delta_{w,j-1} w_{t+1} + \delta_{w,j} w_t + \cdots =$$
$$= w_{t+j} + \delta_{w,1} w_{t+j-1} + \cdots + \delta_{w,j-2}(\varphi_0 + \varepsilon_{t+2} + \theta_1\varepsilon_{t+1} + \theta_2\varepsilon_t) + \delta_{w,j-1}(\varphi_0 + \varepsilon_{t+1} + \theta_1\varepsilon_t + \theta_2\varepsilon_{t-1}) +$$
$$+ \delta_{w,j}(\varphi_0 + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}) + \cdots$$

so $$\frac{\partial Y_{t+j}}{\partial \varepsilon_t} = \delta_{w,j-2}\theta_2 + \delta_{w,j-1}\theta_1 + \delta_{w,j}$$

Exercise 2

The Norwegian GDP data, discussed in the lectures, have been put on the course webpage in a zip-file. The file contains a Stata dta-file (with "tsset" data). The data are also given as an excel-file for those that use another software.

The data are quarterly observations of ln(GDP) in the period 1978q1 – 2013q2.
In this exercise we want to calculate out-of-sample forecasts for the period 2013q3 – 2015q4, based on the Arima(1,1,2) model estimated in the lecture notes 3 (LN3). In addition we want to evaluate the uncertainty of the forecast – using Stata's simulation procedure.

A. (i) Describe formulas for the 4 first forecasts ($\hat{Y}_{t+j|t}$, for $j = 1,2,3,4$ and $t = 2013q2$) (see, e.g., LN4 page 9).
(ii) How would you predict the necessary ε_s's involved?

ANSWER:
(i) From the argument in LN4 page 9-10 we get

$$\hat{Y}_{t+1|t} = \varphi_0 + \varphi_1 Y_t + \theta_1\hat{\varepsilon}_t + \theta_2\hat{\varepsilon}_{t-1}$$
$$\hat{Y}_{t+2|t} = \varphi_0 + \varphi_1 \hat{Y}_{t+1|t} + \theta_2\hat{\varepsilon}_t$$
$$\hat{Y}_{t+j|t} = \varphi_0 + \varphi_1 \hat{Y}_{t+j-1|t} \text{ for } j = 3, 4, \ldots$$

(Printing mistake in the formula in line 5 on page 10 (LN4): the last $\hat{\varepsilon}$ in the formula should be $\hat{\varepsilon}_{t-q}$ (not $\hat{\varepsilon}_{t+j-q}$))

(ii) If the process is invertible, we can express
$$\varepsilon_t = c_0 + Y_t + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \cdots + \gamma_t Y_0 + \cdots \approx c_0 + Y_t + \gamma_1 Y_{t-1} + \gamma_2 Y_{t-2} + \cdots + \gamma_t Y_0$$
truncated at $t = 0$. When t is not too small, this is a good approximation (i.e., when γ_t is negligible).

Note. Apart from this quick and easy approximation, there are developed general and exact recursion formulas for prediction of ε_t in the literature (see, e.g., the "innovation algorithm" in Brockwell & Davis, "Time Series: Theory and

Methods”, Springer 1987 or later, combined with exercise 4 below. See also Hamilton chap. 4)

B. Out-of-sample forecast of ln GDP

Download the GDP data in “Data for seminar III”. Note that the data in the Stata *dta*-file already are declared a time series (by the `tsset` command). The variable “ln BNP” is called `y` in the data set, while Y_t in model (1) is actually Δy (sorry about slightly clumsy use of symbols…)

As a check you may make plots `tsline y` for ln *BNP* and `tsline D.y` for $\Delta \ln BNP$.

Estimate the arima(1,1,2) for *y* by

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

ARIMA regression

Sample:  1978q2 - 2013q2                        Number of obs      =       141
                                                Wald chi2(3)       =     96.22
Log likelihood =  456.0505                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
         D.y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
y            |
       _cons |   .0062495   .0012077     5.17   0.000     .0038826    .0086165
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .7294206   .1237281     5.90   0.000     .4869179    .9719232
             |
          ma |
         L1. |  -1.104681   .1158547    -9.54   0.000    -1.331752   -.8776096
         L2. |   .4904761   .0843566     5.81   0.000     .3251403     .655812
-------------+----------------------------------------------------------------
      /sigma |    .009512   .0006189    15.37   0.000      .008299     .010725
------------------------------------------------------------------------------
```

Considering the small estimated error standard deviation, it may be a good idea (with regard to the simulations below) to “blow up” the observations, e.g., by a factor 100, $yy = 100 * y$,

```
. gen yy=100*y


. arima yy,arima(1,1,2)
ARIMA regression

Sample:  1978q2 - 2013q2                        Number of obs      =       141
                                                Wald chi2(3)       =     96.26
Log likelihood = -193.2779                      Prob > chi2        =    0.0000

------------------------------------------------------------------------------
             |                 OPG
        D.yy |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
yy           |
       _cons |   .6250144   .1207667     5.18   0.000      .388316    .8617127
-------------+----------------------------------------------------------------
ARMA         |
          ar |
         L1. |   .7294721   .1237054     5.90   0.000      .487014    .9719302
             |
          ma |
         L1. |  -1.104718   .1158342    -9.54   0.000    -1.331749   -.8776874
         L2. |   .4904786   .0843513     5.81   0.000      .325153    .6558042
-------------+----------------------------------------------------------------
      /sigma |   .9511711    .061883    15.37   0.000     .8298826     1.07246
------------------------------------------------------------------------------
```

Which estimates in the output are affected by the "blow up" and which are not?

(We see that $\varphi_1, \theta_1, \theta_2$ are not affected, while all the other parameters are affected by a factor 100)

We need to define our forecast horizon (10 quarters ahead)

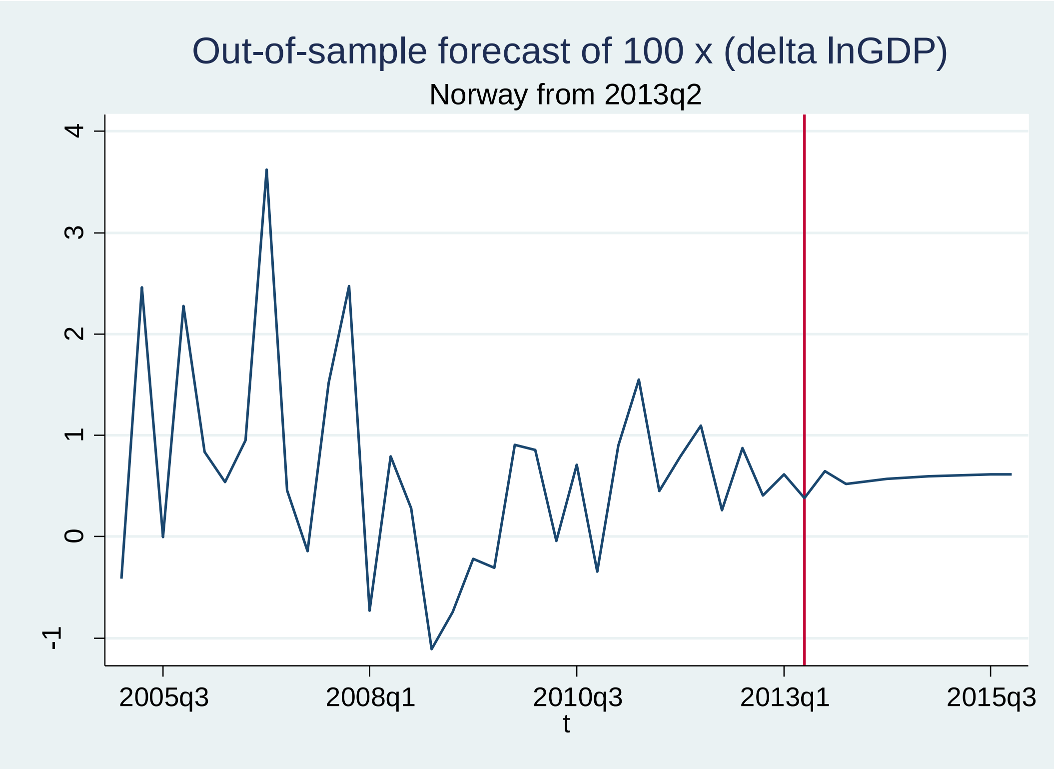
```
. tsappend, add(10)
```

Look at the effect in the data base.
Following the descriptions in the first forecast section in the pdf-manual (read the example of the Klein model there), we need to store the estimates of the arima (this will be used by the forecast commands)

```
. estimates store uts                  (uts is the name of the estimates set)

. estimates dir                        (for checking)
-------------------------------------------------------
        name | command      depvar       npar  title
-------------+-----------------------------------------
         uts | arima        D.yy            5
-------------------------------------------------------

. forecast create mod1, replace        (initiation of the model called mod1)
(Forecast model mod1 ended.)
  Forecast model mod1 started.

. forecast estimates uts, names(uts)   (the forecasts are now ready to calculated)
Added estimation results from arima.
  Forecast model mod1 now contains 1 endogenous variable.

. forecast solve, log(off)
Computing dynamic forecasts for model mod1.
-------------------------------------------
Starting period:  2013q3
Ending period:    2015q4
Forecast prefix:  f_
```

Look at the result in the data base.
Plot the time series $100\Delta \ln BNP$ including the forecasts from 2005q1.

```
. twoway (tsline f_uts if t>=tq(2005q1)) , tline(2013q2)
```



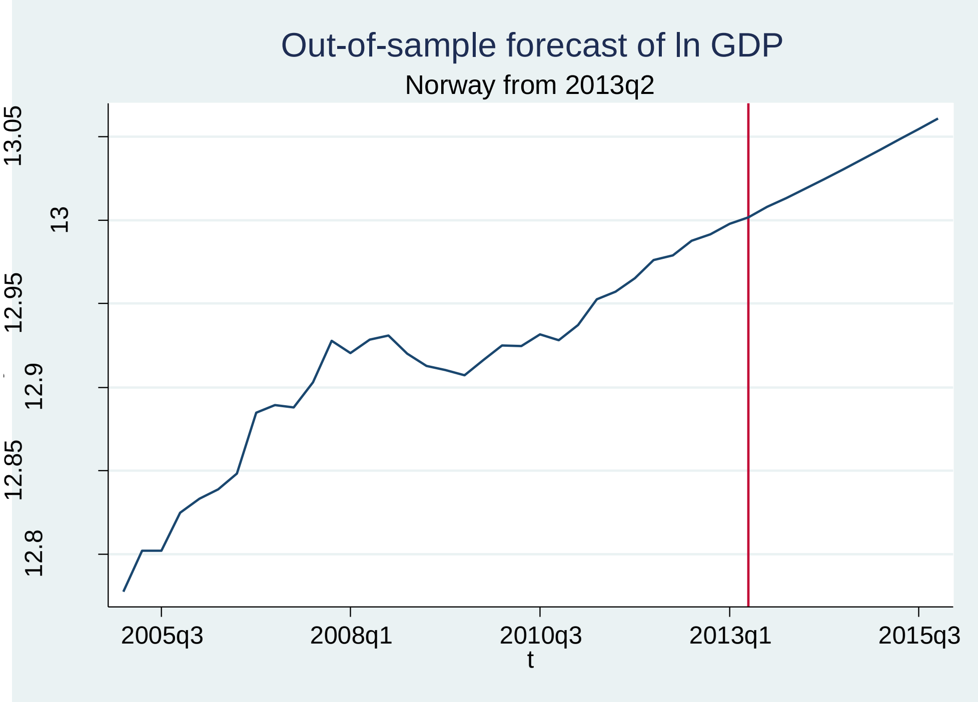
Now, transfer these forecasts to the original series $\Delta \ln BNP$ (in the original scale) and plot the result from 2005q1

[**Hint:** Note that if we have data, $x_0, \Delta x_1, \Delta x_2, \ldots, \Delta x_t$, the we get

$x_t = x_0 + \Delta x_1 + \Delta x_2 + \cdots + \Delta x_t$ (called "integration" of the $\{\Delta x_t\}$ series)]

```
. gen ff_uts=f_uts/100
(1 missing value generated)

. replace ff_uts = 12.1191 in 1
(1 real change made)

. gen yhat=sum( ff_uts)       (The "sum" command generates cumulative sums in Stata)
```

```
. twoway (tsline yhat if t>=tq(2005q1)) , tline(2013q2)
```



Uncertainty by simulation

(In Stata13 read pdf-manual time series –> forecast solve -> example 2 and 3. In particular read the introduction before example 2 on prediction uncertainty simulation.)

We will concentrate on the yy –series ($100\,\Delta \ln BNP$).

Given a model there are two components in the prediction uncertainty, (1) uncertainty due to uncertain estimates in the model, and (2) uncertainty due to variations of the error term.

(**Note.** Of course, there are other sources of uncertainty as well, for example, model specification uncertainty, the risk of structural breaks in the future of the series, etc. These outer sources of uncertainty are ignored in the present calculations…)

For the following simulations, it is a good idea to set the seed for the random number generator before each simulation experiment: (1) That makes it possible to reproduce exactly the experiment, and (2) it is always good to compare the results from several repetitions of the experiment for stability.

First simulate the estimation uncertainty

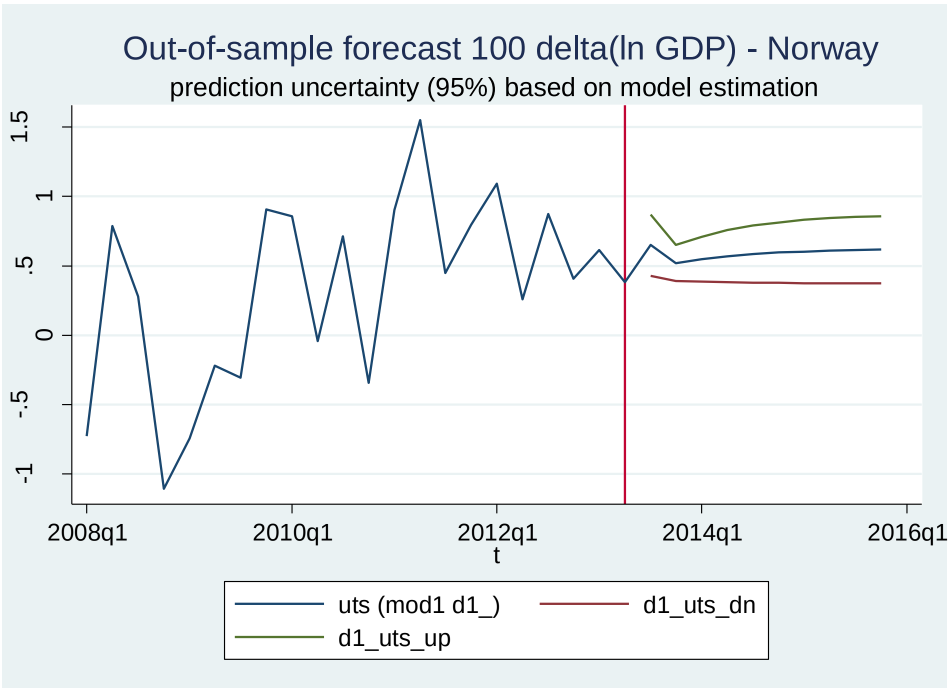
```
. set seed 54321                  (or choose a seed (number) of your own)

. forecast solve, prefix(d1_) begin(tq(2013q3)) simulate(betas, statistic(stddev,
prefix(sd1_)) reps(100))
```

Note. Trying out several seeds (remember to drop the forecast-variables in the data base before each attempt), you will probably discover a source of un-stability of the simulations (not discussed in the Stata manual as far as I know). Sometimes you will even get quite wild (!) results. **Question:** Can you think of any reason (speculation here not knowing exactly how Stata operates its simulations in a dynamic setting) why this phenomenon might occur in this particular situation?

> (My guess, not knowing exactly how Stata operates here, is that the causal stationary solution has a root $\hat{\varphi}_1 = 0.729$ relatively close to the unit circle – implying that draws from the asymptotic normal distribution of the parameter estimators sometimes produce explosive non-stationary solutions that may produce huge error estimates. This is not commented in the Stata manual.)

After having made up your mind of your seed, calculate confidence limits (check the data base for the results)

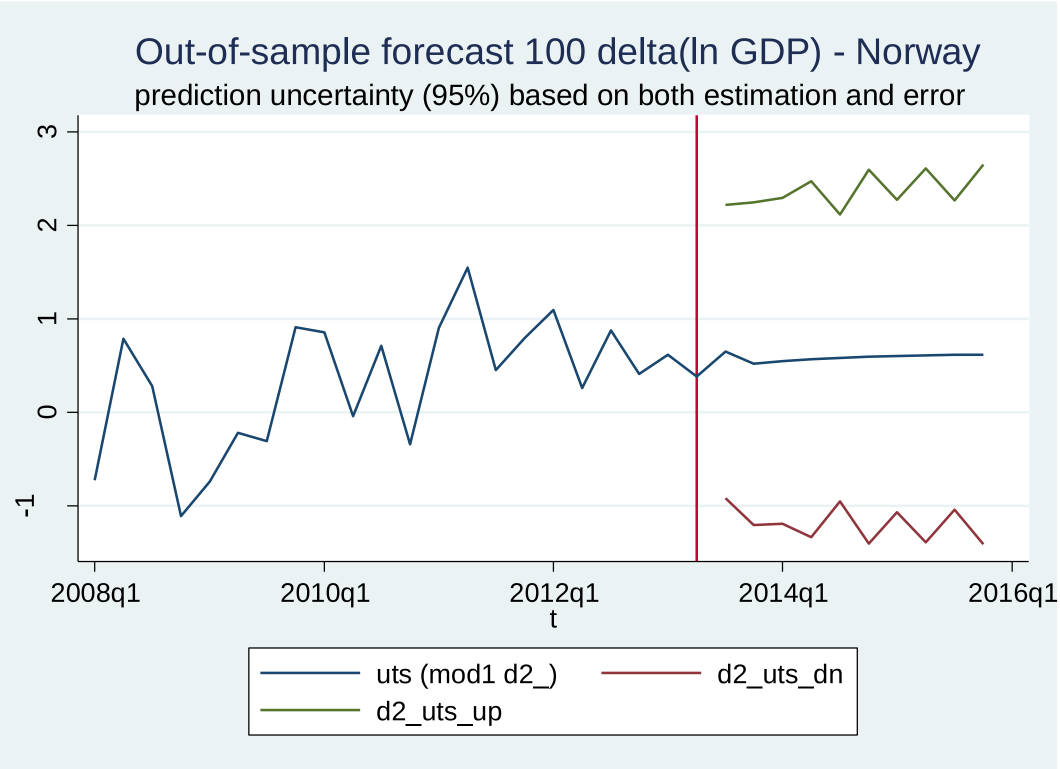
```
. gen d1_uts_up = d1_uts + invnormal(0.975)*sd1_uts
(142 missing values generated)


. gen d1_uts_dn = d1_uts + invnormal(0.025)*sd1_uts
(142 missing values generated)

. twoway (tsline d1_uts if t>=tq(2008q1)) (tsline d1_uts_dn if t>=tq(2008q1))
(tsline d1_uts_up if t>=tq(2008q1)), tline(2013q2)
```



Second, simulate the total (estimation plus error) prediction un-certainty.

```
. set seed 567              (or a seed of your own)

. forecast solve, prefix(d2_) begin(tq(2013q3)) simulate(betas errors,
statistic(stddev,prefix(sd2_)) reps(100)) log(off)

. gen d2_uts_up = d2_uts + invnormal(0.975)*sd2_uts
```

```
. gen d2_uts_dn = d2_uts + invnormal(0.025)*sd2_uts

twoway (tsline d2_uts if t>=tq(2008q1)) (tsline d2_uts_dn if t>=tq(2008q1)) (tsline
d2_uts_up if t>=tq(2008q1)), tline(2013q2)
```



Exercise 3 (Some elementary facts about eigenvalues)

It is important to know about eigenvalues in time series analysis. They are also important in most branches of multivariate analysis.

Read first the review of eigenvalues in LN4 page 14.

(i) Explain why a $k \times k$ matrix A has at most k eigenvalues.

[**Hint.** Consider the meaning of the determinant, $|A - \lambda I|$]

ANSWER:

The determinant $\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk}-\lambda \end{vmatrix}$ is a polynomial of order k. The highest order, k, is obtained from the product along the main diagonal.

(ii) Show that the eigenvalues of a diagonal matrix are its diagonal elements.

Note that a diagonal matrix $A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{kk} \end{pmatrix}$ has determinant $|A| = a_{11}a_{22}\cdots a_{kk}$.

ANSWER: We have

$$|A-\lambda I| = \begin{vmatrix} a_{11}-\lambda & 0 & \cdots & 0 \\ 0 & a_{22}-\lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{kk}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{kk}-\lambda)$$

which is 0 with k solutions, $\lambda = a_{ii}, \quad i = 1, 2, \ldots, k$.

(iii) Let the eigenvalues of the $k \times k$ matrix A be $\lambda_1, \lambda_2, \ldots, \lambda_k$. Show that the determinant is $|A| = \lambda_1\lambda_2\cdots\lambda_k$.

(**Note** (1**)** that this is true also if some of the λ's are equal, which you do not have to prove here.

Note (2) that this implies that A is nonsingular if and only if all eigenvalues are $\neq 0$.

Note (3). It may also be mentioned (that you do not have to prove) that the *rank* of a square matrix(relevant in cointegration analysis) is equal to the number of non-zero eigenvalues.)

> [**Hint.** Use property (19) in LN4 page 15, together with the general matrix property $|C\,D| = |C||D|$ for square matrices of same dimension.]

ANSWER: From the review in LN4, $A = B\Lambda B^{-1}$, where $\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{pmatrix}$

Then $|A| = |B\Lambda B^{-1}| = |B||\Lambda||B^{-1}|$, and

$BB^{-1} = I \quad \Rightarrow \quad 1 = |I| = |BB^{-1}| = |B||B^{-1}| \quad \Rightarrow \quad |B^{-1}| = |B|^{-1}$

Hence $|A| = |\Lambda| = \lambda_1\lambda_2\cdots\lambda_k$

(iv) Show that all eigenvalues of a symmetric square matrix (i.e., $A = A'$) must be real numbers.

(**Note.** Hence, eigenvalues of covariance matrices are always real! They must also be non-negative since, otherwise, we could find random variables with negative variances. A square matrix with only real and non-negative eigenvalues is called *positive semi definite* . If all the eigenvalues are strictly positive, we call the matrix *positive definite*)

[**Hint.** Show first that the standard representation of a complex number, $z = a + ib$, is *unique*, (where *a* is called *the real part* of z and the real number *b the imaginary part* of *z*).

(**Hint:** Suppose $z = a + ib = a' + ib'$ are two representations of *z*. To show that $a = a'$ and $b = b'$, look at the absolute value (modulus) of the difference between the two representations.)

Answer: $0 = a + ib - (a' + ib') = (a - a') + i(b - b')$ with absolute value

$$0 = |0| = |(a - a') + i(b - b')| = \sqrt{(a - a')^2 + (b - b')^2} \quad \Rightarrow \quad a = a' \text{ and } b = b'$$

Now, fill in the details in the following sketch of an argument. You will need the following general matrix rule for taking the transpose of an arbitrary product of matrices, $(CD)' = D'C'$ where $C \sim p \times q$ and $D \sim q \times p$.

Proof (with answer). Let b, λ be any eigenvector and corresponding eigenvalue λ for *A* (i.e., such that $Ab = \lambda b$). Write *b* and λ in standard form, $b = \underline{x} + i\underline{y}$ and $\lambda = c + id$, where $\underline{x}, \underline{y}$ are real vectors and c, d real scalars. We must prove that $d = 0$:

We get

$$Ab = A\underline{x} + iA\underline{y} = \lambda b = (c + id)(\underline{x} + i\underline{y}) = (c\underline{x} - d\underline{y}) + i(c\underline{y} + d\underline{x})$$

Equating the real and imaginary parts of the result, we must have

$A\underline{x} = c\underline{x} - d\underline{y}, \quad A\underline{y} = d\underline{x} + c\underline{y}$.

Then multiply these two equations by $\underline{y}'$ and $\underline{x}'$ respectively (giving $\underline{y}'A\underline{x} = c\underline{y}'\underline{x} - d\underline{y}'\underline{y}$ and $\underline{x}'A\underline{y} = d\underline{x}'\underline{x} + c\underline{x}'\underline{y}$). Now $\underline{y}'A\underline{x}$ and $\underline{x}'A\underline{y}$ must be equal since the transposed of a scalar is the same scalar (i.e., $5' = 5$). Therefore

$\underline{y}'A\underline{x} = (\underline{y}'A\underline{x})' = \underline{x}'A'\underline{y}'' = \underline{x}'A'\underline{y} \overset{A'=A}{=} \underline{x}'A\underline{y}$. Then, noticing $\underline{x}'\underline{y} = (\underline{x}'\underline{y})' = \underline{y}'\underline{x}'' = \underline{y}'\underline{x}$, and subtracting, $0 = \underline{y}'A\underline{x} - \underline{x}'A\underline{y} = c\underline{y}'\underline{x} - d\underline{y}'\underline{y} - (d\underline{x}'\underline{x} + c\underline{x}'\underline{y}) = -d(\underline{x}'\underline{x} + \underline{y}'\underline{y})$ shows that *d* must be 0 since $\underline{x}'\underline{x} + \underline{y}'\underline{y}$ must be $\neq 0$. (**End of proof**).]

A little bit about triangular matrices that are very common (and useful) in multivariate analysis.

(v) Explain why the determinant of an upper (or a lower) triangular matrix must be the product of its main diagonal elements.

Hint. Evaluate, e.g., the determinant along the first column etc.
(An *upper triangular matrix* is

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ 0 & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{kk} \end{pmatrix}$$ (i.e., all elements below main diagonal are 0)

In a *lower triangular matrix* all elements are above the main diagonal are 0.

ANSWER: . Evaluating the determinant along the first column gives

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ 0 & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{kk} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & a_{2k} \\ \vdots & \ddots & \vdots \\ 0 & & a_{kk} \end{vmatrix}$$, where the last determinant is again upper triangular. Repeating this k times, we get $|A| = a_{11}a_{22}\cdots a_{kk}$.

(vi) Explain why the eigenvalues of an upper (or lower) triangular matrix are the elements on the main diagonal.

ANSWER:

$$|A-\lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1k} \\ 0 & a_{22}-\lambda & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{kk}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{kk}-\lambda)$$

Exercize 4 (residuals)

Let $\{Y_t\}$ be any stationary series that admits a linear (causal) representation

$$Y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi \varepsilon_{t-2} + \cdots$$

where $\varepsilon_t \sim WN(0, \sigma^2)$, and where all parameters are known (as estimates or or in other ways). Show that

(*) $\qquad \varepsilon_t = Y_t - \hat{Y}_{t|t-1}$

where $\hat{Y}_{t|t-1}$ is the one-step-ahead predictor of Y_t based on $\{Y_{t-1}, Y_{t-2}, Y_{t-3}, \ldots\}$.

Note. Using (*) we have a natural way to predict the error terms (thus determining residuals) in causal stationary processes (including ARMA(p,q)), based on the one-step-ahead forecasts of the series.

ANSWER: We have $Y_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi \varepsilon_{t-2} + \cdots$, and
$Y_{t|\mathrm{t}-1} = \mu + \psi_1 \varepsilon_{t-1} + \psi \varepsilon_{t-2} + \cdots$
Hence $Y_t - \hat{Y}_{t|t-1} = \varepsilon_t$.