Exam in: ECON 4160: Econometrics: Modelling and Systems Estimation

Day of exam: 11 January 2024 (postponed exam)

Time of day: 09:00—14:00

This is a 5 hour home exam.

Guidelines:
In the grading, question A gets 50 %, B 50 %.

Question A (50 %)

Assume that the time series Y_t is generated by the difference equation:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t, t = 1, 2, ..., T, \quad (1)$$

where ϵ_t denotes an error-term which is gaussian white-noise and linearly uncorrelated with Y_{t-1} and Y_{t-2}. In this question you can take as given that Y_t is a weakly stationary time series.

1. Explain in words why the OLS estimators of the parameters of (1) are consistent under the assumptions given.

   Table 1 contains estimation results for model (1), including a battery of standard residual mis-specification tests and results of dynamic analysis of the estimated equation. The data set is artificial (computer generated) and the model equation corresponds to the DGP.

2. Explain briefly why the results of the mis-specification test are as one would expect, given the information above.

3. Where in the output do you find the relevant information about the stationarity of Y_t?

4. Show that the equation:

   $$\Delta Y_t = \phi_0 + \phi_1^{\dagger} \Delta Y_{t-1} + \gamma Y_{t-1} + \epsilon_t, t = 1, 2, ..., T. \quad (2)$$

   is a re-parameterization of (1) and explain why table 2 is an example of the re-parameterization.

5. Explain why the following equation:

   $$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \Delta \epsilon_t, t = 1, 2, ..., T, : \quad (3)$$

   is *not* a re-parameterization of (1), and explain why table 3 does not provide reliable estimates of the parameters ϕ_1 and ϕ_2.

6. Returning to model (1) and the estimated model equation in table 1.

   (a) What are the optimal forecasts of Y_{101}, Y_{102} and Y_{103}, conditional on $Y_{100} = -0.5$ and $Y_{99} = 1.5$?

   (b) Assume that the forecast horizon is increased towards infinity. Is it then true that the optimal forecast will approach $E(Y) = 0.105$ asymptotically and in a cyclical manner? Explain your answer.

Question B (50 %)

The empirical relationship between aggregate private consumption expenditure and disposable income for households is important to include in an explanatory macroeconomic model for Norway. In the following quarterly time series for these two variables are used.

CP is used to denote consumption while INC denotes income. Both variables are deflated by the Norwegian consumer price index, hence they are measured in real terms (millon 2021-kroner to be precise).

Existing research has shown that the explanation of aggregate consumption is improved if a measure of households' wealth is included. We denote it by W. Also this variable is measured in millon 2021-kroner.

In the following, the natural logarithms of the variables are denoted LCP, LINC and LW. You may take as granted that all three variables are $I(1)$-series.

1. Table 4 contains results for the Johansen-method to cointegration analysis. The results were based on the estimation of a VAR with the three endogenous variables: LCP, LINC and LW. The VAR was specified with fifth order dynamics, and it was not mis-specified.

   Explain why the output supports that the number of cointegration vectors (r) can be set to $r = 1$.

2. Table 5 shows the estimated cointegration parameters (beta in the output), and the associated vector with adjustment coefficients (alpha in the output).

   (a) Do you find the beta vector to be economically meaningful?

   (b) A potential comment to the estimated alpha-coefficients is that they indicate that the Johansen-method was a better test of cointegration than the ECM-test in this case. Do you agree?

3. Conditional on cointegration, an Unrestricted Reduced Form (URF) system can be formulated. The system has DLCP, DLINC and DLW (the "D" denotes difference as usual) as endogenous variables, and contains four lags of each of them as explanatory variables. In addition, each row of the URF includes an intercept and the first lag of the variable:

$$ECM = LCP - 0.64LINC - 0.18LW.$$

   Explain why it is correct to say that the URF only includes time series that are I(0).

4. The log-likelihood of the estimated URF was 1119.42 (using the 1987(1)-2019(4) sample). Table 6 shows estimation results for a model of the URF. The model, which implies 22 zero restrictions on the coefficients of the URF-system, has log-likelihood: 1101.46837.
   Calculate the LR-test of the joint validity of the restrictions, and show that the null hypothesis of joint validity of the restrictions is not rejected at the 5 % significance level.

5. Explain why the equation-by-equation estimation by OLS used in table 6, referred to as "1SLS" in the output, can result in inconsistent estimation of (in particular) the coefficient of DLINC in the equation for DLCP, unless the error term in that equation is uncorrelated with the error-term in the equation for DLINC.

6. Table 7 contains IV-estimation results for the equation for DLCP.

   (a) Based on a comparison of results of the OLS and IV estimations of the coefficient of DLINC in the equation for DLCP, is there a serious simultaneous equations bias here?

   (b) Give an explanation and interpretation of the *Specification test* in table 7.

7. The variable DLINC is said to be weakly exogenous for the parameters of the conditional model equation for DLCP, if those parameters can be efficiently estimated without taking the information of a marginal equation for DLINC into account.

   Explain how you could implement a test of weak exogeneity of DLINC, i.e., if we take the equation for DLCP in table 6 as the conditional model equation in question.

Tables with estimation results

Table 1: Estimation results for model equation (1).

```
Modelling Y by OLS
The estimation sample is: 2 - 100

                 Coefficient  Std.Error  t-value  t-prob
Y_1                  1.13312    0.04717     24.0  0.0000
Y_2                -0.885791    0.04698    -18.9  0.0000
Constant           0.0788564    0.03378     2.33  0.0217

sigma               0.333387  RSS            10.6700882
R^2                 0.863739  F(2,96) =      304.3 [0.000]**
Adj.R^2               0.8609  log-likelihood   -30.2050
no. of observations       99  no. of parameters     3

AR 1-2 test:      F(2,94)   =  0.48809 [0.6153]
ARCH 1-1 test:    F(1,97)   =  0.51255 [0.4758]
Normality test:   Chi^2(2)  =   5.7687 [0.0559]
Hetero test:      F(4,94)   =  0.89091 [0.4726]
Hetero-X test:    F(5,93)   =  0.70860 [0.6184]

Dynamic analysis:

Roots of Y lag polynomial:
        real         imag      modulus
     0.56656      0.75153      0.94116
     0.56656     -0.75153      0.94116
```

Table 2: Estimation results for model equation (2). D has been used to denote the difference operator.

```
Modelling DY by OLS
The estimation sample is: 2 - 100

                 Coefficient  Std.Error  t-value  t-prob
DY_1                0.885791    0.04698     18.9  0.0000
Constant           0.0788564    0.03378     2.33  0.0217
Y_1                -0.752675    0.04223    -17.8  0.0000

sigma               0.333387  RSS              10.6700882
R^2                 0.829344  F(2,96) =      233.3 [0.000]**
Adj.R^2             0.825789  log-likelihood       -30.2050
no. of observations       99  no. of parameters           3


AR 1-2 test:      F(2,94)   =  0.48809 [0.6153]
ARCH 1-1 test:    F(1,97)   =  0.51255 [0.4758]
Normality test:   Chi^2(2)  =   5.7687 [0.0559]
Hetero test:      F(4,94)   =  0.85900 [0.4916]
Hetero-X test:    F(5,93)   =  0.70860 [0.6184]
```

Table 3: Estimation results for model equation (3).

```
 Modelling DY by OLS
 The estimation sample is: 2 - 100

                 Coefficient  Std.Error  t-value  t-prob
DY_1                0.913176    0.06344     14.4  0.0000
DY_2               -0.787767    0.06412    -12.3  0.0000
Constant         0.000710244    0.04337   0.0164  0.9870

sigma               0.431506  RSS             17.8749203
R^2                 0.714111  F(2,96) =      119.9 [0.000]**
Adj.R^2             0.708155  log-likelihood    -55.7447
no. of observations       99  no. of parameters        3


AR 1-2 test:       F(2,94)   =   9.2415 [0.0002]**
ARCH 1-1 test:     F(1,97)   =  0.81518 [0.3688]
Normality test:    Chi^2(2)  =   2.4943 [0.2873]
Hetero test:       F(4,94)   =  0.23142 [0.9201]
Hetero-X test:     F(5,93)   =  0.20023 [0.9616]
```

Table 4: Johansen method: Cointegration rank test results.

```
I(1) cointegration analysis, 1987(1) - 2019(4)

 eigenvalue    loglik for rank
                 1016.515   0
    0.18678      1030.161   1
   0.066977      1034.736   2
 0.00037679      1034.761   3

 H0:rank<=  Trace test [ Prob]
     0          36.492 [0.007] **
     1          9.2007 [0.354]
     2        0.049746 [0.824]
```

Table 5: Johansen method: Estimated $\boldsymbol{\beta}$-vector (conditional on $r = 1$) and the associated $\boldsymbol{\alpha}$-vector.

```
beta
LCP          1.0000
LINC       -0.64116
LW         -0.17566

alpha
LCP        -0.28049
LINC        0.30084
LW         -0.17766

Standard errors of alpha
LCP        0.094130
LINC        0.11570
LW          0.11183
```

Table 6: Estimation results for multiple-equation modelling of DLCP, DLINC and DLW.

```
        Estimating the model by 1SLS
        The estimation sample is: 1987(1) - 2019(4)

 Equation for: DLCP
                  Coefficient  Std.Error  t-value  t-prob
 DLINC               0.227151    0.05339     4.25  0.0000
 DLCP_1             -0.525354    0.09545    -5.50  0.0000
 DLCP_2             -0.460868    0.09007    -5.12  0.0000
 DLCP_3             -0.218328    0.08145    -2.68  0.0084
 DLCP_4              0.174699    0.07033     2.48  0.0143
 DLW_1               0.208474    0.06369     3.27  0.0014
 DLW_2               0.155149    0.06934     2.24  0.0271
 Constant            0.341242     0.1441     2.37  0.0194
 ECM_1              -0.177531    0.07649    -2.32  0.0219

 sigma = 0.0127559

 Equation for: DLW
                  Coefficient  Std.Error  t-value  t-prob
 DLW_1               0.387106    0.08937     4.33  0.0000
 DLW_2               0.144311    0.08675     1.66  0.0988
 Constant            0.332949     0.1723     1.93  0.0556
 ECM_1              -0.173465    0.09120    -1.90  0.0595

 sigma = 0.0199255

 Equation for: DLINC
                  Coefficient  Std.Error  t-value  t-prob
 DLINC_1            -0.504046    0.08291    -6.08  0.0000
 DLINC_2            -0.355243    0.08999    -3.95  0.0001
 DLINC_3            -0.396332    0.08802    -4.50  0.0000
 DLINC_4             0.349818    0.08424     4.15  0.0001
 DLW_2               0.164868    0.05811     2.84  0.0053
 Constant           -0.191994     0.1317    -1.46  0.1475
 ECM_1               0.107391    0.06968     1.54  0.1259


 sigma = 0.0147306

 log-likelihood     1101.46837  -T/2log|Omega|     1663.36803
 no. of observations       132  no. of parameters          29

 LR test of over-identifying restrictions: Chi^2(22) =   35.905 [0.0311]*
```

Table 7: 2SLS estimation results for the equation for DLCP.

```
        Modelling DLCP by IVE
        The estimation sample is: 1987(1) - 2019(4)

                 Coefficient  Std.Error  t-value  t-prob
DLINC         Y     0.255349    0.07515     3.40  0.0009
DLCP_1             -0.506543     0.1019    -4.97  0.0000
DLCP_2             -0.449941    0.09247    -4.87  0.0000
DLCP_3             -0.206915    0.08430    -2.45  0.0155
DLCP_4              0.172485    0.07053     2.45  0.0159
DLW_1               0.204574    0.06418     3.19  0.0018
DLW_2               0.144292    0.07234     1.99  0.0483
Constant            0.373969     0.1567     2.39  0.0186
ECM_1              -0.195018    0.08329    -2.34  0.0209

sigma              0.0129293  RSS            0.020060166
Reduced-form sigma  0.013329
no. endogenous variables   2  no. of instruments       14
no. of observations      132  no. of parameters        12
mean(DLCP)         0.0061532  se(DLCP)          0.0502657
Additional instruments:
[0] = DLINC_1
[1] = DLINC_2
[2] = DLINC_4

Specification test: Chi^2(2)  =  0.91472 [0.6330]
```