Exam in: ECON 4160: Econometrics: Modelling and Systems Estimation—Annotated version

Day of exam: 26 November 2018

Time of day: 09:00—12:00

This is a 3 hour school exam.

Guidelines:
In the grading, question A gets 33 %, B 33 % and C 33 %.

Question A (1/3)

1. Use the information in Table 1 to decide the order of integration of the two equity price indeces LPA (Norway) and $LPAW$ (world), both measured in natural logarithms.

   A: In the two tables for LPA and $LPAW$, all the ADF test statistics indicate non-rejection of the H_0 that the variables are $I(1)$. None of the tests are significant at the 10 % level or lower. If we should specify two ADF-test to use (if the answer depended on that) it would be the ones in the D-lag 3 row for LPA and D-lag 2 row for $LPAW$. This choice is motivated by the p-values in the t-DY_lag column.

   In the two tables for $DLPA$ and $DLPAW$ we also get clear results: rejecting the H_0 that both variables are $I(1)$. The best ADFs to use would be the ones in D-lag 2 row for $DLPA$ and D-lag 1 row for $DLPAW$.

2. In this question you can take for granted that $LPAW_t$ is a strongly exogenous variable in the conditional model for $DLPA_t$ shown in Table 2.

   (a) Based on the mis-specification tests reported, is there any indication that statistical inference based on the *t-values* will be unreliable?

   A: The four tests reported (residual autocorrelation, departure from normality, and two forms of heteroskedasticity) do not indicate *mis-specification (ie no threat to "internal validity"), and on that basis use of the p-values of t-values are relieable.*

(b) Assume that the test situation is:

H_0: No relationship between $DLPA_t$ and $DLPAW_t$ against H_1: There is a relationship between $DLPA_t$ and $DLPAW_t$.

Based on the information in the table, what is your conclusion?

A: The t-value of $DLPAW_t$ of 9.67 implies rejection of H_0 at practically any significance level.

(c) A business school student says that the column labelled *t-prob* (which contains the *p-values*) supports that there is a long-run relationship between LPA_t and $LPAW_t$, because the *p-values* of both LPA_{t-1} and $LPAW_{t-1}$ show that the coefficients are significantly different from zero when a 10 % significance level is used.

Explain why this test method can lead to a spurious relationship.

A: Since the two variables are I(1), the conventional p-values are not valid to use as test statistics. Specifically, the use of a p-value of 0.10 for LPA_{t-1} will lead to much higher Type-I error probability, implying a danger of concluding with a relationship when in fact there is no relationship (spurious relationship)

(d) Explain why the ECM-test is a valid test of the H_0 of absence of cointegration between LPA_t and $LPAW_t$, and why that test will not reject H_0 at the 10 % level of significance (see Table 2 for critical calues).

A: The ECM test compares the t-valus with the correct critical value of the t-values under the null on no relationship between the two $I(1)$ variables. These critical values are located to the left of the corresponding critcal values of the normal or Student t-distribution, because of the "Dickey-Fuller tupe distribtrui" . A valid test is therefore to compare the t-ratio of the coefficient of LPA_{t-1} with a critical value of the distribution of the ECM-test of no long-run relationship. The 10 % critical value given in the table is -2.91, so that the H_0 is not rejected.

(e) Assume that the outcome of the test of the H_0 of absence of cointegration between LPA_t and $LPAW_t$ was different: that the H_0 had been rejected by the ECM-test. In this case, what would the estimated long-run elasticity of PA_t with respect to PAW_t be?

A: The long run elasticity is derived from the stationary solution. It becomes:

$$\text{Long-run elasticity: } \frac{0.0432321}{0.0355075} = 1.2175$$

Question B (1/3)

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

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t, \tag{1}$$

where ε_t is a Gaussian white noise variable with variance σ^2, hence $\varepsilon_t \sim N(0, \sigma^2)$ for all t.

1. Under which condition on ϕ_1 can the stable solution for Y_t be written in terms of a past value Y_{t-1-j} (initial condition), and the white noise terms: $\varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_{t-j}$?

   A: $-1 < \phi_1 < 1$

2. Under the stability condition in QB1, derive the solution for Y_t when $j = 2$.

   A:

$$\begin{aligned} Y_t &= \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t = \phi_0(1 + \phi_1) + \phi_1^2 Y_{t-2} + \varepsilon_t + \phi_1 \varepsilon_{t-1} \\ &= \phi_0(1 + \phi_1 + \phi_1^2) + \phi_1^3 Y_{t-3} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} \end{aligned}$$

3. Under the stability condition in QB1, what is the expression for the solution for Y_t when $j \to \infty$?

   A:

$$Y_t = \frac{\phi_0}{1 - \phi_1} + \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{t-j}$$

4. Assume that we are interested in estimating the parameters ϕ_0 and ϕ_1.

- Explain why Y_{t-1} is a pre-determined variable in (1).

  A:The solution for Y_{t-1} is:

$$Y_{t-1} = \frac{\phi_0}{1-\phi_1} + \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{t-1-j}$$

  showing that Y_{t-1} is dependent on ε_{t-1}, ε_{t-2},ε_{t-3},.... Hence by definition, Y_{t-1} is pre-determined.

- Denote the OLS estimator by $\hat{\phi}_1$. Explain why $\hat{\phi}_1$ is a biased estimator in any finite sample, but that $plim(\hat{\phi}_1 - \phi_1) = 0$ under the stability condition.

  A:The OLS estimator can be written as:

$$\hat{\phi}_1 = \phi_1 + \frac{\sum_t \varepsilon_t (Y_{t-1} - \bar{Y})}{\sum_t (Y_{t-1} - \bar{Y})^2}$$

$$E(\hat{\phi}_1 - \phi_1) = E\left(\frac{\sum_t \varepsilon_t (Y_{t-1} - \bar{Y})}{\sum_t (Y_{t-1} - \bar{Y})^2}\right)$$

  We cannot show formally that the right-hand side is zero. In the case of $0 < \phi_1 < 1$, intuitively, there will be a tendency towards negative correlation for a given sample in the numerator, since $Y_{t-1} - \bar{Y}$ is relatively smooth. However, as $t \to \infty$ the proportion of independent information increases so that $plim(\hat{\phi}_1 - \phi_1) = 0$, intuitivly.

5. Assume that we are interested in forecasting Y_{T+h}, using (1). For simplicity, we assume that ϕ_0, ϕ_1, σ^2 and Y_T are known numbers. Since we do not know the future white noise variables, the forecast is made by replacing ε_{T+1},ε_{T+2}, ..., ε_{T+h} by zeros (expected values). Denote the sequence of (point) forecasts by Y^f_{T+h}, $h = 1, 2, .., H$.

   Show that:

$$Y^f_{T+h|T} \to \frac{\phi_0}{1-\phi_1} \text{ as } h \to \infty, \tag{2}$$

   under the stability condition.

A: There are several ways to show this. One is to make use of the solution of the difference equation conditional on Y_T:

$$Y_{T+1} = \phi_0 + \phi_1 Y_T + \varepsilon_{T+1}$$
$$Y_{T+2} = \phi_0 + \phi_1 Y_{T+1} + \varepsilon_{T+2} = \phi_0(1+\phi_1) + \phi_1^2 Y_T + \varepsilon_{T+2} + \phi_1 \varepsilon_{T+1}$$
$$....$$
$$Y_{T+h} = \phi_0(1+\phi_1 + .. + \phi_1^{h-1}) + \phi_1^h Y_T + \sum_{j=0}^{h-1} \phi_1^j \varepsilon_{T+h-j}$$

Define

$$Y^f_{T+h|T} = E(Y_{T+h} \mid Y_T)$$

hence

$$Y^f_{T+h|T} = \phi_0(1+\phi_1 + .. + \phi_1^{h-1}) + \phi_1^h Y_T$$

from the solution (and the information in the question). It follows directly from this that in the stationary case:

$$Y^f_{T+h} \rightarrow \frac{\phi_0}{1-\phi_1} \text{ as } h \rightarrow \infty.$$

Some students may deliver another, equivalent expressions, like:

$$Y^f_{T+h} = \frac{\phi_0}{1-\phi_1} + \phi_1^h(Y_T - \frac{\phi_0}{1-\phi_1}), \; -1 < \phi_1 < 1$$

where we have used that

$$1 + \phi_1 + .. + \phi_1^{h-1} = \frac{1-\phi_1^h}{1-\phi_1},$$

from the formula for the sum of the h first terms in a geometric progression.

6. Assume that right after you published your forecast, there is a structural break so that (1) changes to:

$$Y_{t+h} = (\phi_0 + d) + \phi_1 Y_{t+h-1} + \varepsilon_{t+h}, \text{ for } t = T \text{ and } h = 1, 2, ..., H$$

(a) Assume that $d > 0$. How will the expected forecast-error $(Y_{T+1} - Y^f_{T+1|T})$ be affected by this structural break?

A: The forecast is:

$$Y^f_{T+1|T} = \phi_0 + \phi_1 Y_T$$

The economy goes:

$$Y_{T+1} = (\phi_0 + d) + \phi_1 Y_T + \varepsilon_{T+1}$$

$$(Y_{T+1} - Y^f_{T+1|T}) = d + \varepsilon_{T+1} > 0$$

(b) Will (2) hold in this case?

A: Yes (2) holds: The forecasts equilbrium corrects to the pre-break expectation of Y_t.

The most "direct" interpretation is that d represents a permanent change in the expectation of the process. In that case there is a bias in the forecast-error $(Y_{T+h} - Y^f_{T+h|T})$ also for large h. *However, if the break is temporary, so that for example $d = 0$ again after one period,* $(Y_{T+h} - Y^f_{T+h|T}) \to 0$ as $h \to \infty$.

Question C (1/3)

Consider the macro model:

(3) $$C_t = c_0 + c_1 GDP_t + c_2 C_{t-1} + \epsilon_{Ct}$$

(4) $$J_t = d_0 + d_1 GDP_t + d_2 GDP_{t-1} + d_3 J_{t-1} + \epsilon_{Jt}$$

(5) $$GDP_t = C_t + J_t + G_t$$

The endogenous variables are: C_t (private consumption), GDP_t (gross domestic product), J_t (private investment). G_t (public expenditure) is determined outside the system, it is an exogenous variable.

Assume that the coefficients different from zero. Assume that the two error-terms are Gaussian white noise variables. We assume that the covariance matrix of the error terms (ie $\mathbf{\Omega}$) is invertible, but it is not necessarily a diagonal matrix.

1. What are the conditions for stationarity of the time series variable GDP_t is this model? (No derivations are required in the answer).

   A: All eigenvalues of the associated companion form matrix must be less than one in magnitude.

2. Explain why the OLS estimator of c_1 is an inconsistent estimator (simultaneity bias)

   A: From the reduced form: GDP_t is correlated with ϵ_{Ct}. Therefore the OLS estimator is not consistent. Since the inconsistency is due to the simultaneity is is called simultaneity bias.

3. Are (3) and (4) identified on the order condition?

   A: (3) is identified on the order condition. The n-1 rule requires 2, while the equation omits J_t, J_{t-1}, GDP_{t-1} and G_t. (4) *also identified, but with lower order of identification.*

4. Explain why 2SLS is a more efficient method of estimation than IV for an over-identified structural equation.

   A: *2SLS (also called GIVE) makes use of optimal instruments. They are symptotically uncorrelated with the disturbance term and have a higher correlation the included endogenous variables than any individual instrumental variable. The optimal instruments are the predicted values of the endogenous variables obtained from the reduced form.*

5. Imagine that you have been able to estimate the SEM (3)-(5) by FIML. Will the estimates be identical to the 2SLS estimates? Explain.

   A: FIML takes account of (and estimates) the off-diagonal elements in Ω. Hence in general, FIML gives different estimates than 2SLS. The expection is when the equations are exectly idenified, or if Ω is the identify matrix.

6. Assume also that you are interested in testing the hypothesis H_0: $d_2 = -d_1$ (the restriction implying that J_t depends on ΔGDP_t). Explain in words how you can test this hypothesis by the use of FIML or 2SLS estimation.

   A: Estimate without the restriction and with the restriction. Use the log-likelihood test. Another test method is to reparameterize equation

(4)

$$J_t = d_0 + d_1 \Delta GDP_t + (d_1 + d_2) GDP_{t-1} + d_3 J_{t-1} + \epsilon_{Jt}$$

and use the "t-value" for 2SLS/FIML estimated coeffeficient of GDP_t *to test the hypothesis.*

Tables

```
Unit-root tests
The dataset is: C:\SW20\ECON4160\H2018\Exam\MODobligexam.in7
The sample is: 1973(2) - 2018(2) (185 observations and 4 variables)

LPA: ADF tests (T=181, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 3      -1.093      0.99148   0.1065      2.481  0.0140
 2     -0.9773      0.99228   0.1080     -1.636  0.1037
 1      -1.083      0.99142   0.1085      4.241  0.0000
 0     -0.8944      0.99259   0.1135

LPAW: ADF tests (T=181, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 3     -0.6086      0.99715  0.05602      1.564  0.1197
 2     -0.5399      0.99746  0.05625     -2.180  0.0306
 1     -0.6487      0.99692  0.05684      5.691  0.0000
 0     -0.3780      0.99806  0.06162

DLPA: ADF tests (T=181, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 3      -6.544**    0.25367   0.1056      2.025  0.0444
 2      -6.222**    0.35127   0.1065     -2.436  0.0158
 1      -8.997**    0.20990   0.1080      1.705  0.0900
 0      -9.868**    0.29873   0.1086

DLPAW: ADF tests (T=181, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 3      -6.093**    0.34188  0.05601     0.6527  0.5148
 2      -6.466**    0.37264  0.05592     -1.542  0.1249
 1      -8.687**    0.29101  0.05613      2.216  0.0280
 0      -8.853**    0.39075  0.05674
```

Table 1: Test results for LPA_t, the natural logarithm of the Oslo Stock Exchange Index, and LPAW, the log of a world equity price index.

```
EQ(1)  Modelling DLPA by OLS
      The dataset is: C:\SW20\ECON4160\H2018\Exam\MODobligexam.in7
      The estimation sample is: 1985(1) - 2018(2)

                  Coefficient  Std.Error  t-value  t-prob
DLPA_1               0.118291    0.04598     2.57  0.0113
LPA_1              -0.0355075    0.01952    -1.82  0.0713
LPAW_1              0.0432321    0.02383     1.81  0.0721
DLPAW                0.899811    0.09305     9.67  0.0000
DLSPOILUSD           0.153493    0.03232     4.75  0.0000
DVOLUSA           -0.00389039   0.001015    -3.83  0.0002
DPADUM               0.975029     0.1063     9.17  0.0000
Constant           0.00718436   0.007280    0.987  0.3256

sigma               0.0491035  RSS             0.303805505
R^2                  0.799467  F(7,126) =    71.76 [0.000]**
Adj.R^2              0.788327  log-likelihood       217.839
no. of observations       134  no. of parameters          8

AR 1-5 test:      F(5,121)  =   1.3873  [0.2338]
ARCH 1-4 test:    F(4,126)  =   0.68241 [0.6054]
Normality test:   Chi^2(2)  =   1.5245  [0.4666]
Hetero-X test:    F(34,99)  =   1.3095  [0.1536]

Critical values of ECM-test: -3.21 (5 %), -2.91 (10 %)
```

Table 2: Estimation results for a model of DLPA which is the first difference of LPA_t. DLPAW is the first difference of LPAW. The other variables are the change in the log of the oil price ($DLSPOILUSD$) , a measure of volatily ($DVOLUSA$) and a dummy ($DPADUM$).