Exam in: ECON 4160: Econometrics: Modelling and Systems Estimation. Version with annotations for evaluators.

Day of exam: 1 December 2022

Time of day: 09:00—14:00

This is a 5 hour home exam.

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

Question A (33 %)

1. Consider the cobweb model with Q_t (quantity) and P_t (price) as endogenous variables. Show that the autoregressive coefficient in the final form equation for P_t is the ratio between the slope coefficients of supply (c) and demand (a):

$$\phi_1 = \frac{c}{a}, \tag{1}$$

where we have used ϕ_1 to denote the autoregressive coefficient in the final form equation for P_t.
A: The cob-web model has been a red-thread in the course. For this question is does not matter whether the deterministic or stochastic version of the model is used. The deterministic version that we have used in teaching and in the book has been:

$$\begin{aligned} Q_t &= aP_t + b_t,\ a < 0\,,\ \text{demand} \\ Q_t &= cP_{t-1} + d, b > 0,\ \text{supply} \end{aligned}$$

The term and concept of final form equation has also been central in the course and in the curriculum.

The final form of equation for P_t in the cobweb-model: Re-normalize the demand-equation

$$P_t = \frac{1}{a}Q_t - \frac{1}{a}b_t$$

and substitute Q_t by the right hand side of the supply equation to obtain the final form equation of P_t as:

$$\begin{aligned} P_t &= \frac{c}{a}P_{t-1} + \frac{d}{a} - \frac{b_t}{a} \\ &= \phi_1 P_{t-1} + \frac{d - b_t}{a} \end{aligned}$$

where ϕ_1 represents the notation we have used for the first autoregressive coefficient.

2. Consider more specifically the stochastic version of the cobweb model, with white-noise random disturbances in the two equations. What is the condition for weak stationary of the time series P_t and Q_t which are generated by the respective final equations?
A: The final form equation for P_t is a first order stochastic difference equation. The condition for stationarity:

$$\mid \phi_1 \mid \neq 1$$

If we restrict the data generation to the casual solution, the condition is:

$$-1 < \phi_1 < 1$$

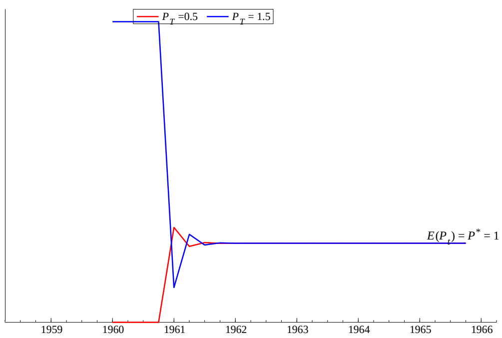
Given that the casual solution has been used 99 % of the time, $-1 < \phi_1 < 1$ is a good enough answer. Since the homogenous parts of the final form equations of the two endogenous variables are identical (general result),the same condition implies stationarity of Q_t. Intuitively: The two endogenous variables of the model cannot have their separate stability/stationarity properties.

3. Consider the particular special case where the final form equation for P_t is:

$$P_t = -0.2P_{t-1} + 1.2 + \varepsilon_{Pt}, t = 0, 1, ..., T, \tag{2}$$

where ε_{Pt} is white-noise. Assume, abstracting from estimation, that you are given the task of forecasting P_t for the period $t = T+1, T+2, ..., T+H$ where H denotes the forecast horizon.

(a) Draw (rough) sketches of the forecast function $E(Y_{T+h} \mid \mathcal{I}_T)$, $h = 1, 2, ..., H$ for two different initial conditions: $P_T = 0.5$ and $P_T = 1.5$.
HINT: Since the absolute value of the autocorrelation coefficient is relatively small, setting H to for example 5 may be enough to get a clear picture.
A: A rough picture (simulated, since I cannot easily do a real drawn sketch in the way the student will do) could be:



(b) What does the two forecasts have in common, and why?
A: The two forecasts have in common that they have the same (asymptotic) "endpoint", which is identical to the unconditional expectation. This is due to the stationarity of the series which implies that forecasts are equilibrium correcting ("glide-path" analogy). The oscillations are also common, and are due to the negative autoregressive coefficient.

4. Assume that in period $t = T + 1$, ie, in the first period after you have produced your forecast, there is a structural break in the economy (ie., in the DGP) so that the final form equation becomes:

$$P_t = -0.2P_{t-1} + 2.4 + \varepsilon_{Pt}, t = T + 1, ..., T + H, \tag{3}$$

with no change in the distribution of ε_{Pt}.
Explain how this break in the DGP will affect the forecast errors of the dynamic forecasts $E(P_{T+h} \mid \mathcal{I}_T)$, $h = 1, 2, ..., H$, for the case where the initialization price was $P_T = 0.5$.
A: The forecast errors (defined as $P_{T+h} - E(P_{T+h} \mid \mathcal{I}_T)$) have positive biases. After the break $E(P_t) = 2$ while the forecast still "believe that it is 1.

5. Imagine that in period $T + 1$, you are asked to update your forecast. Assume that you can observe $P_{T+1} = 2$, but that you do not yet know the coefficient values in the new final form equation (3). Can you suggest ways of robustifying the forecasts for $t = T + 2, ..., T + H$ which are conditional on $P_{T+1} = 2$?

A: Intercept correction is one possibility; In practice can add the observed forecast error for period $T+1$ to the forecasting function used for the updated forecast. Since the structural break is in the long-run mean of P_t, that will reduce the forecast-error compared to do correction.

Some possible details: Conditional on $P_{T+1} = 2$ the economy evolves according to the new DGP:

$$P_{T+2} = -0.2 * 2 + 2.4 + \varepsilon_{PT+2} = 2 + \varepsilon_{PT+2}$$

while the ' 'raw" forecast from the model is:

$$P^f_{T+2} = -0.2 * 2 + 1.2 = 0.8$$

so on average ($\varepsilon_{PT+2} \approx 0$) will have a forecast-error of:

$$P_{T+2} - P^f_{T+2} = 2 - 0.8 = 1.2$$

The forecast error you observed for $T + 1$ was:

$$e_{T+1} = 2 - (-0.2 * 0.5 + 1.2) = 0.9$$

The intercept-corrected forecast $T + 2$ is then:

$$\begin{aligned} \tilde{P}^f_{T+2} &= -0.2 * 2 + 1.2 + e_{T+1} \\ &= -0.2 * 2 + 1.2 + 0.9 = 1.7 \end{aligned}$$

and the forecast-error of the so called robustified forecast (set ($\varepsilon_{PT+2} \approx 0$)) will be

$$P_{T+2} - \tilde{P}^f_{T+2} = 2 + \varepsilon_{PT+2} - 1.7 = 0.3$$

which is already much better than the "raw" error 1.2.

Another kind of robustification is to increase the forecast confidence region, to allow for the "more uncertainty" that a forecasted may induce form observing the first large forecast error. This may be the only option in the case where we do not know yet that $T + 1$ error case caused by a permanent break. (Other forms of robust forecast, like "over-differencing" as an kind of implied intercept-correction are relevant, if mentioned, but it was not time to demonstrate them in class).

Question B (33 %)

Consider, as a generalization of the cobweb model, the partial equilibrium model of Q_t and P_t with X_{1t} and X_{2t} as strictly exogenous explanatory variables:

$$Q_t + a_{12,0}P_t = a_{10} + a_{11,1}P_{t-1} + a_{12,1}Q_{t-1} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + \epsilon_{Qt}, \quad (4)$$

$$a_{21,0}Q_t + P_t = a_{20} + a_{21,1}P_{t-1} + a_{22,1}Q_{t-1} + \gamma_{21}X_{1t} + \gamma_{22}X_{2t} + \epsilon_{Pt}, \quad (5)$$

To complete the econometric model specification, it is assumed that the error-terms ϵ_{Qt} and ϵ_{Pt} are jointly normally distributed, conditional on Q_{t-1}, P_{t-1}, X_{1t} and X_{2t}. The variances of the error-terms are denoted ω^2_Q and ω^2_P, and the covariance is denoted ω_{PQ}.

1. Give an example of a set of parameter restrictions that would define a recursive model, as a special case of the SEM we have specified.
   A: For example $a_{21,0} = 0$ and $\omega_{PQ} = 0$.

2. Assume that the simultaneous equations (4)-(5) can be solved (ie., we do *not* consider a recursive model). The reduced form model is a VAR. Denote the variances and the covariance of the VAR error-terms by σ_Q^2, σ_P^2, and σ_{PQ}. The VAR has an equivalent statistical representation which consists of a conditional model equation and a marginal model equation. Describe the structure of those two model equations for the case where we are interested in the conditional model of Q_t given P_t.
   NOTE: With the exception of the coefficient of P_t, call it β_1, you are not expected to give algebraic expressions for the parameters of the conditional and marginal model equations. It is enough to explain verbally which variables are on the right hand side of each of the two model equations.
   A: On the RHS of the conditional model for Q_t we have P_t with coefficient $\beta_1 = \frac{\sigma_{PQ}}{\sigma_P^2}$,and $P_{t-1}, Q_{t-1}, X_{1t}, X_{2t}$, constant and error-term. On the RHS of the marginal equation we have $P_{t-1}, Q_{t-1}, X_{1t}, X_{2t}$ constant and error-term-

3. P_t is an endogenous explanatory variable in the conditional model equation. Does this imply that the OLS estimator of β_1 is inconsistent? Explain you answer.
   A: No, by valid conditioning P_t is uncorrelated with the error-term in the conditional model equation.

4. We now return to the SEM we specified at the start of this question and discuss identification. In this question it assumed that the variances and the covariance of the SEM error-terms are unrestricted. Discuss identification in the following situations:

   (a) $\gamma_{12} = \gamma_{22} = 0$; all other coefficients non-zero. A: (4) and (5) not identified (order and rank apply)

   (b) $\gamma_{12} = 0$; all other coefficients non-zero. A: (4) identified, (5) is not identified,

   (c) $a_{12,1} = \gamma_{11} = \gamma_{22} = 0$; all other coefficients non-zero. A: (4) over-identified, (5) is identified,

5. It is custom to say about the 2SLS estimator that it makes use of "optimal instrumental variables". Can you give an explanation of the meaning of that statement?
   A: The statement is relevant for over-id model equations. The optimality of the instruments stems from the first stage, where the reduced form coefficients of the instrumental variables are estimated by minimization of the sum of squared residuals, so the weights attached to each individual IV is optimal in that meaning.

Question C (33 %)

We have a quarterly temperature time series, TEMP_t.[1] In line with standard notation, the quarterly difference is denoted DTEMP_t, ie., $\text{DTEMP}_t = \text{TEMP}_1 - \text{TEMP}_{t-1}$.

1. Explain how the information in Table 1 can be used to conclude, based on the use of formal statistical tests, that $\text{TEMP}_t \sim I(1)$.
   A: Important to use a ADF statistic that is based on a not-misspecified Dicky-Fuller regression, as far as it is possible to see from the output. So, rejecting the null hypothesis for TEMP_t using the ADF for D-lag 0 is a wrong answer. The use of the row with D-lag 4 is a good answer, and leads to not rejecting. The table for DTEMP_t is simpler since all ADF-statistics reject. But formally, the rows for D-lag 3 and D-lag 4 give correct ADF-test to use.

2. We also have a quarterly time series of CO_2 in the atmosphere on Mauna Loa (Hawaii) in the northern Pacific Ocean. Denote the natural logarithm of this variable by $LCO2$. Assume that $LCO2_t \sim I(1)$. Explain how the information in Table 2 can be used

[1] Western-hemisphere land-ocean temperature in degrees Celsius as a deviation from the mean temperature in the period 1951-1980.

to conclude, based on the use of formal statistical tests, that TEMP_t and LCO2_t are cointegrated.
A: This is an (unrestricted) conditional model in ECM-form. The ECM test and the critical values from the Ericsson and MacKinnan paper can be used. The value of the t_{ECM} test is -5.01 which is significant. It supports the validity of the test, that the test statistic for residual autocorrelation is insignificant.

3. Explain why the implied ECM variable is:

$$\text{ecmTEMPLCO2}_t = \text{TEMP}_t - 4.8\text{LCO2}_t, \tag{6}$$

where the cointegration parameter is given with one decimal point.
A:

$$\frac{1.57595}{0.329043} = 4.7895 \approx 4.8$$

4. Table 3 shows the estimation results for the model of $DTEMP_t$ conditional on $TEMP_t$ and $LCO2_t$ being cointegrated, while Table 4 shows the results for a marginal equation for $DLCO2$.

   (a) Logically speaking, why cannot both TEMP_t and LCO2_t be weakly exogenous with respect to the cointegration parameter?
   A: Granger's representation theorem: Cointegration implies equilibrium correction in at least one of variable in the cointegrated system, hence both variables cannot be weakly exogenous.

   (b) How can you use Table 3 and Table 4 to assess the possible weak exogeneity of one of the variables DTEMP_t and DLCO2_t with respect to the cointegration parameter, and what is your conclusion?
   A: Already the significance of the t-value of ecmTEMLCO2_{t-1} in Table 3 is evidence that DTEMP_t is not weakly exogenous. Furthermore, as the ECM-variable is insignificant in Table 4, we have evidence that DLCO2_t is weakly exogenous, while DTEMP_t is not weakly exogenous. This is a good enough answer at this level.
   However, there is a weakness in that argument as well, since the definition of weak exogeneity is with respect to the marginal equations of the variable in question. And we do not have the marginal equation for DTEMP_t here! However, an indirect argument is possible. First, we have the marginal equation for DLCO2_t with an estimated marginal equilibrium correction coefficient, call it τ_2. The null hypothesis of $\tau_2 = 0$ cannot be rejected in Table 4. As regression (conditional modelling) theory tells us, the adjustment coefficient in Table 3 is in principle:

$$\alpha = \tau_1 - \kappa\tau_2$$

   where τ_1 is the equilibrium correction coefficient in the marginal equation for DTEMP_t, and κ is the regression coefficient in the model of DTEMP_t conditional on DLCO2_t. Now, as $\tau_2 = 0$, it follows that the significant ECM term in Table 3 is indeed also evidence of an (underlying) significant equilibrium correction coefficient in the marginal equation for DTEMP_t.

5. Table 5 shows results obtained by using the Johansen-method for cointegration analysis. Are these results (by and large) in support of the results obtained earlier in this question?
A: First, a remark that this method does not assume anything about weak exogeneity with respect to cointegration parameters is a good remark. Then to the test: The test of cointegration in Table 5 is the Trace test. It rejects no-cointegration, rank =0, against one relationship, rank =1 (two would imply that the two variables were $I(0)$, an internal logical inconsistency). Hence, the Trace-test is supportive of how we

concluded by using the ECM-test. Since we have rank=1 and one single relationship, the vector of cointegration parameters is identifed (subject to normalization on one of the variables). The estimated ECM-variable from the Johansen method is:

$$\text{JohaecmTEMPLCO2}_t = \text{TEMP}_t - 5.9\text{LCO2}_t \tag{7}$$

so that the Johansen-estimated cointegration coefficient is larger in magnitude than that of the ECM estimate. The difference is not trivial, but not so large that there is any contradiction here. We are looking at finite sample results from two different estimators.

Finally there is direct evidence in Table 5 in support of weak WE of LCO2_t, as the standard error of the estimated alpha of LCO2 is so large that the marginal equilibrium correction coefficient is not significantly different from zero.

Tables with estimation results and facimile of table with critical values for ECM-test

Table 1: Dickey-Fuller tests of unit-root. TEMP_t and DTEMP_t.

```
The sample is: 1959(2) - 2022(3) (254 observations)

TEMP: ADF tests (T=251, Constant; 5%=-2.87 1%=-3.46)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 4     -1.207       0.97079   0.1622      1.012  0.3126
 3     -1.092       0.97378   0.1623     -2.968  0.0033
 2     -1.511       0.96352   0.1648     -5.655  0.0000
 1     -2.448       0.93839   0.1748     -2.981  0.0032
 0     -3.092*      0.92270   0.1775

DTEMP: ADF tests (T=251, Constant; 5%=-2.87 1%=-3.46)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 4     -9.104**     -1.0123   0.1623      1.197  0.2325
 3     -10.03**    -0.86981   0.1624    -0.8702  0.3851
 2     -14.40**    -0.97936   0.1623      3.157  0.0018
 1     -17.85**    -0.65434   0.1652      6.021  0.0000
 0     -19.71**    -0.21869   0.1765
```

Table 2: Estimation results for a model of DTEMP_t

```
EQ( 1) Modelling DTEMP by OLS
     The estimation sample is: 1959(2) - 2022(3)

                 Coefficient  Std.Error  t-value  t-prob
DTEMP_1            -0.131801    0.07731    -1.70  0.0895
DTEMP_2            -0.142193    0.07502    -1.90  0.0592
DTEMP_3          -0.00838194    0.06940   -0.121  0.9040
DTEMP_4          -0.00779189    0.06356   -0.123  0.9025
Constant            -8.63916      1.961    -4.41  0.0000
TEMP_1             -0.329043    0.06562    -5.01  0.0000
DLCO2                23.7800      10.16     2.34  0.0201
DLCO2_1              32.8985      10.01     3.29  0.0012
DLCO2_2             -20.7048      10.25    -2.02  0.0446
DLCO2_3              1.40087      10.17    0.138  0.8905
DLCO2_4             -9.45658      10.10   -0.936  0.3502
LCO2_1               1.57595     0.3367     4.68  0.0000
Seasonal           -0.696419     0.1955    -3.56  0.0004
Seasonal_1         -0.932327     0.2572    -3.63  0.0004
Seasonal_2         -0.424001     0.1993    -2.13  0.0344

sigma               0.140492  RSS               4.71737253
R^2                 0.422551  F(14,239) =   12.49 [0.000]**
Adj.R^2             0.388725  log-likelihood       145.822
no. of observations      254  no. of parameters         15
mean(DTEMP)       0.00346457  se(DTEMP)           0.179694

AR 1-5 test:      F(5,234)  =   1.2239 [0.2986]
```

Table 3: Estimation results for a model of DTEMP_t, conditional on cointegration.

```
EQ( 2) Modelling DTEMP by OLS
      The estimation sample is: 1959(2) - 2022(3)

                  Coefficient  Std.Error  t-value  t-prob
DTEMP_1             -0.131776    0.07714    -1.71  0.0889
DTEMP_2             -0.142090    0.07475    -1.90  0.0585
DTEMP_3           -0.00815509    0.06867   -0.119  0.9056
DTEMP_4           -0.00757659    0.06286   -0.121  0.9042
Constant             -8.65646      1.833    -4.72  0.0000
ecmTEMPLCO2_1       -0.328954    0.06539    -5.03  0.0000
DLCO2                 23.7087      9.736     2.44  0.0156
DLCO2_1               32.8042      9.268     3.54  0.0005
DLCO2_2              -20.8142      9.271    -2.25  0.0257
DLCO2_3               1.30000      9.328    0.139  0.8893
DLCO2_4              -9.52581      9.703   -0.982  0.3272
Seasonal            -0.696195     0.1949    -3.57  0.0004
Seasonal_1          -0.932209     0.2566    -3.63  0.0003
Seasonal_2          -0.424476     0.1980    -2.14  0.0330

sigma                0.140199  RSS              4.71738508
R^2                  0.422549  F(13,240) =    13.51 [0.000]**
Adj.R^2              0.391271  log-likelihood      145.822
no. of observations       254  no. of parameters        14
mean(DTEMP)        0.00346457  se(DTEMP)          0.179694

AR 1-5 test:      F(5,235)  =   1.2157 [0.3024]
```

Table 4: Estimation results for a marginal model of $\mathrm{DLCO2}_t$, conditional on cointegration.

```
EQ(3) Modelling DLCO2 by OLS
      The estimation sample is: 1959(2) - 2022(3)

                  Coefficient  Std.Error  t-value  t-prob
DLCO2_1             -0.106905    0.06093    -1.75  0.0806
DLCO2_2             -0.129611    0.06077    -2.13  0.0340
DLCO2_3             0.0494536    0.06163    0.802  0.4231
DLCO2_4              0.304635    0.06112     4.98  0.0000
Constant           -0.0162792    0.01208    -1.35  0.1790
DTEMP_1            0.00151583  0.0005009     3.03  0.0027
DTEMP_2            0.00126003  0.0004879     2.58  0.0104
DTEMP_3            0.00141311  0.0004451     3.17  0.0017
DTEMP_4            0.00125382  0.0004080     3.07  0.0024
ecmTEMPLCO2_1    -0.000536080  0.0004312    -1.24  0.2150
Seasonal           0.00590183   0.001232     4.79  0.0000
Seasonal_1         0.00716308   0.001634     4.38  0.0000
Seasonal_2        -0.00369491   0.001288    -2.87  0.0045

sigma              0.00092758  RSS          0.000207357403
R^2                  0.985325  F(12,241) =     1348 [0.000]**
Adj.R^2              0.984594  log-likelihood      1419.93
no. of observations       254  no. of parameters        13
mean(DLCO2)        0.00109223  se(DLCO2)        0.00747327

AR 1-5 test:      F(5,236)  =   1.0565 [0.3787]
```

Table 5: Results for the Johansen-method for cointegration analysis.

```
I(1) cointegration analysis, 1959(2) - 2022(3)

 eigenvalue    loglik for rank
                 1513.841   0
    0.13991      1532.982   1
  0.0056153      1533.697   2

 H0:rank<=  Trace test [ Prob]
     0         39.712 [0.000] **
     1         1.4303 [0.232]

Asymptotic p-values based on: Unrestricted trend and constant
Unrestricted variables:
Constant
Trend
Number of lags used in the analysis: 5

beta (scaled on diagonal; cointegrating vectors in columns)
TEMP         1.0000    -0.010200
LCO2        -5.896       1.0000

alpha
TEMP       -0.54616     -0.30961
LCO2        0.00023     -0.010983

Standard errors of alpha, conditional om r=1
TEMP       0.088758
LCO2       0.000647
```

Table 6: Facsimile from article by Ericsson and MacKinnon.

304 *Neil R. Ericsson and James G. MacKinnon*

Table 3. Response surface estimates for critical values of the ECM test of cointegration $\kappa_c(k)$: with a constant term.

k	Size (%)	θ_∞	(s.e.)	θ_1	θ_2	θ_3	$\hat{\sigma}$
1	1	−3.4307	(0.0006)	−6.52	−4.7	−10	0.00790
	5	−2.8617	(0.0003)	−2.81	−3.2	37	0.00431
	10	−2.5668	(0.0003)	−1.56	2.1	−29	0.00332
2	1	−3.7948	(0.0006)	−7.87	−3.6	−28	0.00847
	5	−3.2145	(0.0003)	−3.21	−2.0	17	0.00438
	10	−2.9083	(0.0002)	−1.55	1.9	−25	0.00338
3	1	−4.0947	(0.0005)	−8.59	−2.0	−65	0.00857
	5	−3.5057	(0.0003)	−3.27	1.1	−34	0.00462
	10	−3.1924	(0.0002)	−1.23	2.1	−39	0.00364
	1	−4.3555	(0.0006)	8.90	−6.7	−31	0.00959