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

Day of exam: 25 November 2015

Time of day: 09:00—12:00

This is a 3 hour school exam.

Guidelines:
In the grading, question A will count 30%, and question B will count 70%
.

Question A (30 %)

1. Assume that the conditional expectation of Y grows linearly with X. Consider the n variable pairs $(Y_1, X_1), \ldots, (Y_n, X_n)$, and assume that the variable pairs are mutually independent and have identical normal distributions. In this case, what are the expressions for the maximum likelihood estimators of the two parameters

$$\frac{\partial}{\partial X_i} E(Y_i \mid X_i) \text{ and } Var(\varepsilon_i \mid X_i) = \sigma^2 \text{ ?}$$

   Answer: $\frac{\partial}{\partial X_i} E(Y_i \mid X_i)$ is the partial derivative of $E(Y_i \mid X_i)$ which is linear in parameters, and therefore $E(Y_i \mid X_i) = \beta_0 + \beta_1 X_i$ and $\frac{\partial}{\partial X_i} E(Y_i \mid X_i) =: \beta_1$ for example. Based on the assumptions given, the MLE of β_1 is therefore the OLS estimator $\hat{\beta}_1 = \sum_{i=1}^{n}(X_i - \bar{X})Y_i / \sum_{i=1}^{n}(X_i - \bar{X})^2$.

   ε_i is the disturbance term in the model equation $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$. $Var(\varepsilon_i \mid X_i) = \sigma^2$ says that the conditional variance of the disturbance is independent of X (homoscedasticity). Based on the information given, the MLE of σ^2 is therefore $\hat{\sigma}^2 = n^{-1} \sum \hat{\varepsilon}_i^2$ where $\varepsilon_i =: Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$ with OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$.

2. How can you estimate the parameter $\frac{\partial}{\partial X_i} E(Y_i \mid X_i)$ efficiently, if the assumption $Var(\varepsilon_i \mid X_i) = \sigma^2$ is changed to $Var(\varepsilon_i \mid X_i) = \sigma^2 X_i^2$?

Answer: With heteroscedastic error of this particular type, the MLE estimator can be found by estimating the following model equation by OLS

$$\underbrace{\frac{Y_i}{X_i}}_{V_i} = \beta_0 \underbrace{\frac{1}{X_i}}_{W_i} + \beta_1 + \underbrace{\frac{\varepsilon_i}{X_i}}_{e_i}$$

this is because $Var(\frac{\varepsilon_i}{X_i} \mid X_i) = Var(\frac{\varepsilon_i}{X_i} \mid X_i) = \frac{1}{X_i^2} Var(\varepsilon_i \mid X_i) = \sigma^2$, meaning that in the model that regress V_i on W_i and a constant "back" in the homoscedastic model of (1). The efficient estimator of β_1 is therefore the OLS estimator, call it $\tilde{\beta}_1$, of the constant term in the transformed model. This estimator is also known as weighted least squares, which is an example of GLS.

It can be written as (but this is not asked for)

$$\tilde{\beta}_1 = \bar{V} - \tilde{\beta}_0 \bar{W}$$

where $\tilde{\beta}_0$ is the OLS estimator of the slope coefficient in the transformed model.

3. Assume that the relationship

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_{1i}, \tag{1}$$

is an equation in a model consisting of two equations. Discuss identification and estimation of the parameter β_2 in the following three cases: (We denote parameters in the second equation by γ_j in all three cases, and that the only unobservable variables are disturbances.)

(a) The second equation is

$$Z_i = \gamma_0 + \gamma_1 X_i + \varepsilon_{2i} \tag{2}$$

and we assume $Cov(\varepsilon_{1i}, \varepsilon_{2i}) = \omega_{12} \neq 0$, and that X_i is uncorrelated with both disturbances.

Answer: A discussion of identification is based on a classification of variables as endogenous and exogenous. In this case a reasonable classification is the Y and Z are endogenous, while X (as a conditioning variable) is exogenous. Based on that, and since the covariance between the disturbances is unrestricted, (1) and (2) is a "system of regression equations" (it could be the reduced form of a SEM for Y_i and Z_i for example). β_2 is identified in this case and the simple OLS estimator of the parameter is consistent and efficient (since the right hand side variable is the same in both equations, in fact GLS "reduces to" and becomed identical with OLS in this case).

(b) The second equation is

$$X_i = \gamma_0 + \varepsilon_{2i} \tag{3}$$

and we assume $\omega_{12} = 0$.

Answer: In this case, the reasonable classification is that Y and X are endogenous. However, since $\omega_{12} = 0$, the rank and order conditions do not apply. (1) and (3) are equations of a recursive model, and again OLS on (1) gives a consistent estimator of β_1.

(c) The second equation is

$$X_i = \gamma_0 + \gamma_1 Y_i + \gamma_2 Z_{1i} + \gamma_3 Z_{2i} + \varepsilon_{2i} \tag{4}$$

and we assume $\omega_{12} \neq 0$.

Answer: There are two equations in the model and five variables: Y, X, Z_1, Z_2 and *Constant*. We base our answer on the premise that Y and X are endogenous, while Z_1 and Z_2 are exogenous and uncorrelated with both disturbances. Since the correlation between the two disturbances is unrestricted, the rank and order conditions apply. If both $\gamma_2 \neq 0$ and $\gamma_3 \neq 0$, equation (1) is overidentified on the rank condition. 2SLS is then a consistent estimator of β_1 which uses the two instrumental variables in an optimal way (combine them in a way that gives the best linear predictor of X_i). 2SLS is however not statistically efficient since it does not take account of the correlation between the disturbances. Based on an (extra) normality assumption, FIML is more efficient than 2SLS.

4. In case 3(b) and 3(c), is the second equation of the model identified?

Answer: Yes in (b) because of recursiveness. No in (c) (order and rank conditions)

Question B (70 %)

We have collected annual data for hourly wages in Norwegian manufacturing for the period 1970 to 2013. We also have data for the value of labour productivity in this sector. In the print-out from PcGive in Table 1, we denote the logarithms of these two variables as LW (wages) and LZ (value of labour productivity) respectively.

1. Explain why the evidence in Table 1 gives reason to conclude that both LW and LZ are integrated of order one, $I(1)$. (You can consider the degree of augmentation as a given thing).

```
Unit-root tests
The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
The sample is: 1972 - 2013 (44 observations and 2 variables)

LW: ADF tests (T=42, Constant+Trend; 5%=-3.52 1%=-4.19)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
  1     -2.811       0.94888  0.01856      3.623  0.0008
  0     -4.145*      0.92148  0.02125

LZ: ADF tests (T=42, Constant+Trend; 5%=-3.52 1%=-4.19)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
  1     -1.887       0.93151  0.03735    -0.3623  0.7192
  0     -1.915       0.93131  0.03693


Unit-root tests
The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
The sample is: 1973 - 2013 (43 observations and 2 variables)

DLW: ADF tests (T=41, Constant; 5%=-2.93 1%=-3.60)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
  1     -3.079**     0.55138  0.02025      1.143  0.2603
  0     -2.864**     0.62787  0.02033

DLZ: ADF tests (T=41, Constant; 5%=-2.93 1%=-3.60)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
  1     -4.655**   -0.11150  0.03953     0.2770  0.7833
  0     -6.553**   -0.063503  0.03905
```

Table 1: Augmented Dickey-Fuller (ADF) tests to determine the order of integration of LW_t and LZ_t.

Answer: In the upper part of the table the null hypothesis is that LW and LZ are I(1). By using the relevant "t-adf" values (for LW in particular the D-lag 1 row column) we conclude the unit root hypothesis is not rejected for any of the variables.

It is not expected, but would be positive extra if students comment on the role of deterministic augmentation (Constant + trend) as a way of obtaining a well behaved test, for example allowing for a trend in the series both under the null and the alternative.

In the bottom part the null hypothesis is that the differences of LW and LZ are I(1), and this hypothesis is rejected for both. So the overall conclusion is that they are both I(1).

2. According to theory, the system of collective wage bargaining in Norway creates a long-run dependency between the hourly wage and the manufacturing firms' ability to pay, as measured by the value of labour productivity.

   To test this theory, we estimate the following Engle-Granger regression by OLS:

   $$LW_t = \beta_0 + \beta_1 LZ_t + u_t, \qquad t = 1970, ..., 2013 \tag{5}$$

   where u_t is the disturbance. In the print-out in Table 2, the residual from the Engle-Granger regression has been labeled EGLWresiduals.

   Use the results in Table 2 to form a conclusion about whether the theory of a long-run relationship between LW and LZ is supported or not. The critical values of the Engle-Granger test of no long-run (cointegrating) relationship are: 5% $= -3.33$, and 1% $= -3.90$. As part of your answer, explain why you use these critical values, instead of the critical values for the ADF tests given in Table 1.

```
EQ( 1) Modelling LW by OLS
    The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
    The estimation sample is: 1970 - 2013

                Coefficient  Std.Error  t-value  t-prob Part.R^2
Constant          -0.172383    0.04508    -3.82  0.0004   0.2582
LZ                 0.963630   0.008637     112.  0.0000   0.9966

sigma              0.0504044  RSS            0.106705506
R^2                 0.996637  F(1,42) =    1.245e+004 [0.000]**


Unit-root tests
The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
The sample is: 1973 - 2013 (43 observations and 1 variables)

EGLWresiduals: ADF tests (T=41; 5%=-1.95 1%=-2.62)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob       AIC  F-prob
1     -2.853**     0.61462  0.03731     0.5051  0.6163    -6.529
0     -2.935**     0.64303  0.03696                       -6.572  0.6163
```

Table 2: Results for an Engle-Granger regression between LW_t and LZ_t, and unit-root tests for the disturbance of that regression.

Answer: The null hypothesis of no cointegration implies that u_t is I(1) so a relevant test is an ADF test. However in place of the unobservable u_t we use the OLS residuals, and this explains why the critical values are larger in absolute value here than in an ADF for an observable variable. Briefly explained this is because, also under the null of no-relationship, the residuals will "look more" stationary since OLS always finds the best fitting linear relationship in any given sample. Hence, the location of the Enge Granger test of no realtionship is shifted to the left as more regressors are included in the hypothesized cointegratiion regression. Failure to account for this is one aspect of the spurious regression case. In this case, the test does not reject the null hypothesis of no relationship.

3. An alternative test for the null hypothesis of no long-run relationship can be based on a conditional equilibrium correction model (ECM). You find estimation results for such a model in Table 3. In that print-out, the variables DLW and DLZ are the differences of LW and LZ, for example DLW=LW-LW_1, where LW_1 denotes the first lag of LW.

   The ECM of LW also includes two other conditioning variables: DLKPI, which is the inflation rate, and DLNH, which is the change in the length of the normal working day. It is relevant to condition on these two variables because compensation for increases in the cost of living, and for shorter-hours is part of the bargaining between the unions and the firms. We base our analysis on the assumption that DLKPI and DLNH are $I(0)$ variables.

(a) Based on the results in Table 3, and the information that the relevant 1% critical value for the ECM-test of no relationship is -3.29, explain why it is reasonable to conclude that LW is cointegrated with LZ.

Answer: The value of the ECM-test for no cointegration is found as -5.31, which is strongly significant when judged against the 1 % critical value of -3.29.

(b) Can you give some intuition on why the evidence in support of cointegration may be stronger when we use the ECM-test than when the Engle-Granger test is used?

Answer: In the lectures and in the book we learn about the common factor restrictions that are implicit in the Engle-Granger test. This restriction essentially entails that the short-run effect of an increase in LZ is the same as the short-run effect. From the results we see that is far from the case here. A known implication of invalid common factor restrictions is that the power of the Engle-Granger test will be lower than the power of the ECM-test. Another feature worth noting is that DLKPI (in particular), but also DLH are highly significant. Inclusion of these variables may also increase the power of the test.

(c) Use the results to find the estimated long-run elasticity of the wage level (W) with respect to the value of labour productivity (Z). The estimated standard-error of the long-run elasticity can be shown to be 0.032. Is a long-run elasticity of 1 supported empirically if you use a significance level of 5%?

Answer: $0.237388/0.250837 = 0.9464$(with four decimals). The upper limit of an approximate 95 % confidence interval becomes $0.9464 + 2 * 0.032 = 1.0104$, which is somewhat larger than 1. Hence the unit long run elasticity is supported by this constructed confidence interval.

(d) How could the standard-error be calculated? Explain in words.

Answer: The long run elasticity is a ratio of two estimated parameters. The standard-error of that ratio can be calculated by the use of the so called Delta-method (or Bårdsen method). It requires that we have the access to the estimated covariance of the estimated parameters of LW_1 and LZ_1. Alternatively we can re-parameterize the ECM as an ARDL, re-estimate is and obtain the standard error from the menu with Dynamic analysis.

```
EQ(2) Modelling DLW by OLS
      The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
      The estimation sample is: 1972 - 2013

                 Coefficient  Std.Error  t-value  t-prob Part.R^2
DLW_1               0.247629    0.09799     2.53  0.0163   0.1581
Constant         -0.00492564    0.03616   -0.136  0.8925   0.0005
DLZ                 0.100648    0.05653     1.78  0.0839   0.0853
DLZ_1             0.00917294    0.05715    0.160  0.8734   0.0008
LW_1               -0.250837    0.04727    -5.31  0.0000   0.4530
LZ_1                0.237388    0.04547     5.22  0.0000   0.4449
DLKPI               0.567936     0.1122     5.06  0.0000   0.4297
DLNH               -0.685046     0.2384    -2.87  0.0070   0.1953

sigma              0.0124393  RSS              0.00526103907
R^2                 0.911595  F(7,34) =     50.08 [0.000]**
Adj.R^2             0.893394  log-likelihood         129.092
no. of observations       42  no. of parameters            8
mean(DLW)          0.0691441  se(DLW)              0.0380983

AR 1-2 test:      F(2,32)   =   0.55526[0.5792]
ARCH 1-1 test:    F(1,40)   =   1.9135 [0.1743]
Normality test:   Chi^2(2)  =  0.33955 [0.8439]
Hetero test:      F(14,27)  =   1.3089 [0.2652]
RESET23 test:     F(2,32)   =   3.0802 [0.0598]
```

Table 3: Results for ECM of LW_t

4. As a way of imposing a long-run elasticity of one, we define the variable

$$ECMwage = LW - LZ,$$

which we assume to be $I(0)$ from now on, and re-estimate the ECM for wages. Table 4 shows the results. Table 5 shows the result for a marginal model of DLZ.

(a) Explain how you can test the weak exogenity of LZ with respect to the cointegration parameters with the aid of the information given. What does the evidence indicate?

Answer: The relevant statistic to look for here is the t-value of $ECMwage\_1$ in the equation for DLZ, this is with reference to Grangers-representation theorem for example. The evidence suggest rejection of WE, at least at the 5 % level.

(b) Assume that you are asked by the Norwegian Productivity Commission to estimate the dynamic effects on wages of a shock to average labour productivity. Could you use the results reported in Table 4 and 5 to give an answer? Explain how you would motivate your answer.

Answer: To answer the question we now an empirical model of the relationship between productivity and wages. Table 4 and 5 give such a model, for example LZ can be decomposed as $LZ_t = LP_t + LQ_t$ where LP is the log of the producer price index and LQ is log of average labour productivity, so LZ increases one to one with an increase in productivity. One possibility then is to use EQ(3) in Table 4. That the residual misspecification test are insignificant (with a possible exception for the RESET-test which we have not given much attention in the course) is clearly relevant, since this implies that the statistical evidence is reliable. Hence, one relevant answer is to report the dynamic multipliers of LW with respect to a change in LZ that we can calculate from EQ (3) in 4.

However Table 4 and 5 do show evidence of two-way causality between LZ and LW, and that goes in the direction of not (only) relying on the single equation dynamic multipliers form EQ(5), but also the impulse-responses of LW with respect to a shock to the EQ(4) in Table 5.

(c) A colleague suggests that to appropriately address question 4 (b), a SVAR model needs to be considered. Do you think using a SVAR is suitable to deal with this question?

Answer: The answer is "YES" because is important to have identified dynamic multipliers, or impulse responses, and using a SVAR is one coherent way of securing that kind of identification. Would expect that many students introduce the a SEM model for DLW and DLZ, derive the reduced form, and point out that the VAR residuals are correlated. Then introduce recursiveness perhaps. Such answers should count of course!

However, it is possible that some will try to argue that Table 4 and 5 represent a relevant structure. If we for example define $y_1 = DLW$, $y_2 = DLZ, x_1 = ECMwage, x_2 = DLKPI, x_3 = DLH$, we can interpret EQ(3) and EQ(4) as a "conditonal + marginal" model of the VAR-X:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} Const_1 \\ Const_2 \end{pmatrix} + \Phi \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \Upsilon \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{1t} \end{pmatrix}$$

where the matrix Φ is 2×2 and Υ is 2×2. In this interpretation the maximized log likelihood function of the 2-equation model defined by EQ(3) and EQ(4) will be identical to the maximized log likelihood function of VAR-EX- However the disturbances of the conditional model EQ(3) will be uncorrelated with the disturbance of the marginal EQ(3), while ε_{1t} and ε_{2t} in the VAR-EX in general is correlated. This means that we can interpret the impulse response to a shock to EQ(4) as due to shock to LZ, not LW.

Clearly, if the relevant DGP is a different VAR than the one we have just formulated, it does not follow that the disturbances of EQ(3) and EQ(4) are orthogonal.

```
EQ( 3) Modelling DLW by OLS
      The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
      The estimation sample is: 1972 - 2013

                 Coefficient  Std.Error  t-value  t-prob Part.R^2
DLW_1               0.333612    0.09621     3.47  0.0014   0.2557
Constant          -0.0768366    0.01971    -3.90  0.0004   0.3027
DLZ                 0.134225    0.05796     2.32  0.0265   0.1329
DLZ_1              0.0570952    0.05652     1.01  0.3193   0.0283
DLKPI               0.681517     0.1071     6.36  0.0000   0.5365
DLNH               -0.582600     0.2485    -2.34  0.0249   0.1357
ECMwage_1          -0.210894    0.04668    -4.52  0.0001   0.3683

sigma              0.0131938  RSS            0.00609264708
R^2                 0.897621  F(6,35) =     51.14 [0.000]**
Adj.R^2              0.88007  log-likelihood        126.01
no. of observations       42  no. of parameters          7
mean(DLW)          0.0691441  se(DLW)            0.0380983

AR 1-2 test:      F(2,33)   =   2.7736 [0.0770]
ARCH 1-1 test:    F(1,40)   =0.00067955 [0.9793]
Normality test:   Chi^2(2)  =   1.9124 [0.3844]
Hetero test:      F(12,29)  =  0.84777 [0.6042]
Hetero-X test:    F(27,14)  =   1.4910 [0.2185]
RESET23 test:     F(2,33)   =   3.5640 [0.0397]*
```

Table 4: Results for an ECM of LW_t with cointegration imposed in the form of the variable ECMwage.

```
EQ( 4) Modelling DLZ by OLS
      The dataset is: C:\SW20\ECON4160\H2015\Exam\LoennIndogFastland.in7
      The estimation sample is: 1972 - 2013

                 Coefficient  Std.Error  t-value  t-prob Part.R^2
DLZ_1               0.105475     0.1616    0.653  0.5180   0.0117
Constant            0.130999    0.05231     2.50  0.0169   0.1484
DLW_1               0.558570     0.2605     2.14  0.0389   0.1132
DLKPI              -0.144616     0.3070   -0.471  0.6404   0.0061
DLNH               0.0181977     0.7147   0.0255  0.9798   0.0000
ECMwage_1           0.284411     0.1256     2.26  0.0297   0.1247

sigma               0.037941  RSS              0.0518226244
R^2                 0.452038  F(5,36) =      5.94 [0.000]**
Adj.R^2             0.375932  log-likelihood        81.0541
no. of observations       42  no. of parameters           6
mean(DLZ)          0.0693511  se(DLZ)             0.0480278

AR 1-2 test:      F(2,34)   =  0.18983 [0.8280]
ARCH 1-1 test:    F(1,40)   =  0.44294 [0.5095]
Normality test:   Chi^2(2)  =  0.31240 [0.8554]
Hetero test:      F(10,31)  =  0.60697 [0.7958]
Hetero-X test:    F(20,21)  =   1.6381 [0.1349]
RESET23 test:     F(2,34)   =   4.8045 [0.0145]*
```

Table 5: Results for a marginal model of DLZ_t .