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

Day of exam: 28 November 2017

Time of day: 09:00—12:00

This is a 3 hour school exam.

Guidelines:
In the grading, question A gets 20 %, B 40 % and C 40 %.

Question A (20 %)

Consider the linear in parameters system consisting of the four random variables X_i, Y_i, W_i, Z_i.Assume that you have a sample of n mutually independent quadruplets (X_i,Y_i, W_i,Z_i) that are identically distributed. Analyze the difference between the "plim" of the OLS estimate of the total derivative of Y with respect to X and the true partial derivative of Y with respect to X.

Answer note

Omitted variables question. Covered also by obligatory.

Question B (40 %)

A white-noise time series $\{\varepsilon_t; t = 1, 2, ...T\}$ is characterized by:

(1) $$E(\varepsilon_t) = 0$$

(2) $$Var(\varepsilon_t) = \sigma_\varepsilon^2 > 0$$

(3) $$Cov(\varepsilon_t, \varepsilon_{t-j}) = 0, \text{ for } j = \pm 1, \pm 2, ..$$

Assume that the DGP for the time series Y_t is:

(4) $$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

Based on (1)-(4), answer the following questions:

1. What are the expressions for $E(Y_t \mid Y_{t-1})$ and $Var(Y_t \mid Y_{t-1})$?

Answer:

$$E(Y_t \mid Y_{t-1}) = E(\phi_0 + \phi_1 Y_{t-1} + \varepsilon_t) = \phi_0 + \phi_1 Y_{t-1}$$
$$Var(Y_t \mid Y_{t-1}) = Var(\phi_0 + \phi_1 Y_{t-1} + \varepsilon_t) = \sigma_\varepsilon^2$$

2. What are the expressions for $E(Y_{t+1} \mid Y_{t-1})$ and, $Var(Y_{t+1} \mid Y_{t-1})$?

   Answer:

$$\begin{aligned} E(Y_{t+1} \mid Y_{t-1}) &= E(\phi_0 + \phi_1\phi_0 + \phi_1^2 Y_{t-1} + \varepsilon_{t+1} + \phi_1\varepsilon_t) \\ &= \phi_0(1+\phi_1) + \phi_1^2 Y_{t-1} \\ Var(Y_{t+1} \mid Y_{t-1}) &= Var(\phi_0 + \phi_1\phi_0 + \phi_1^2 Y_{t-1} + \varepsilon_{t+1} + \phi_1\varepsilon_t) \\ &= \sigma_\varepsilon^2(1+\phi_1^2) \end{aligned}$$

3. Give the expression for $E(Y_t)$ in the case where Y_t is a covariance stationary time series, denoted $Y_t \sim I(0)$.

   Answer

$$E(Y_t) = \frac{\phi_0}{1-\phi_1}, \text{ when } -1 < \phi_1 < 1$$

   independent of time.

4. Give the expression for $E(Y_t \mid Y_0)$ when Y_t is a unit-root non stationary time series, $I(1)$. Regard Y_0 as a fixed number.

   Answer

   By repeated substitution back to Y_0

$$\begin{aligned} E(Y_t) &= \sum_{i=1}^{t} \phi_0 i + Y_0 = \\ &= \phi_0 t + Y_0,\ t = 1, 2, .. \end{aligned}$$

5. Show that $Var(Y_t)$ is finite if $Y_t \sim I(0)$, but that $Var(Y_t)$ is infinite if $Y_t \sim I(1)$.

   Answer

   By repeated substitution back to Y_0 :

$$Var(Y_t) = \sigma_\varepsilon^2 \sum_{i=1}^{t} \phi_1^2 = \frac{\sigma_\varepsilon^2}{1-\phi_1^2}$$

when $Y_t \sim I(0) \iff -1 < \phi_1 < 1$. But

$$Var(Y_t) \underset{t\to\infty}{\to} \infty$$

when $Y_t \sim I(1) \iff \phi_1 = 1$ or $\phi_1 = -1$.

6. Assume that ε_t is Gaussian white noise.

   (a) Explain in words why the OLS estimator $\hat{\phi}_1$ of ϕ_1 can be interpreted as a Maximum Likelihood estimator.
   Answer

In the case of Gaussian distribution, the log likelihood function of Y_T, $Y_{T-1}, ..., Y_1$ can maximized in two step. The first is to minimize the residual sum of squares:

$$\sum_{i=1}^{T} \left(Y_t - \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} \right)^2$$

which means that the OLS estimators $\hat{\phi}_1$ and $\hat{\phi}_0$ are MLEs

   (a) Explain in words why $\hat{\phi}_1$ is a biased estimator in finite samples, but that it is also a consistent estimator.
   Answer

Y_{t-1} is not strictly exogenous in (4). Hence we know from basic econometrics that $E(\hat{\phi}_1 - \phi_1) \neq 0$ for any finite sample size T. But Y_{t-1} is also pre-determined given the assumption of the statistical model and this implies that

$$\text{plim}(\hat{\phi}_1) = \phi_1 \text{ when } -1 < \phi_1 < 0.$$

($\hat{\phi}_0$ has the same property.)

7. Replace assumptions (1)-(3), by

(5) $$\varepsilon_t = \tau\varepsilon_{t-1} + \nu_t\,, \quad -1 < \tau < 1,$$

where ν_t is white-noise. Will a regression of Y_t on Y_{t-1} give a consistent estimator of ϕ_1 in this case? Explain you answer.
   Answer

No, because in this case Y_{t-1} is no longer pre-determined, due to the AR(1) process in ε_t, Y_{t-1} is correlated with ε_t and ε_{t+1} for example.

8. Under the same assumption as in Q7, can you suggest a method that can be used for consistent estimation of the characteristic roots of the DGP for Y_t?

   Answer: Insert (5) in (4)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \tau\varepsilon_{t-1} + \nu_t$$

   Lag and multiply by τ, subtract the resulting equation from (4) to give the DGP of as

$$Y_t = \phi_0(1-\tau) + (\phi_1 - \tau)Y_{t-1} - \tau\phi_1 Y_{t-2} + \nu_t$$

   where ν_t is white-noise. The coefficients of this AR(2) model equation can be concistently estimated by OLS, and the consistent estimates of the roots can be worked out. Other methods are also possible: Using iteration to obtain consistent estimators of τ and ϕ_1 for example ("RALS", "Cochrane-Orcutt", different names for the same type of method).

Question C (40 %)

Consider the simultaneous equations model (SEM) with two endogenous variables, Y_t and X_t, and two strongly exogenous variables W_t and Z_t.

$$Y_t - \beta_{10}X_t = \beta_{11}X_{t-1} + \phi Y_{t-1} + \beta_{20}Z_t + \epsilon_{1t}, \tag{6}$$

$$-\gamma_{10}Y_t + X_t = \alpha X_{t-1} + \gamma_{11}Y_{t-1} + \gamma_{20}W_t + \epsilon_{2t}. \tag{7}$$

ϵ_{1t} and ϵ_{2t} are white-noise disturbances which may be correlated $Cov(\epsilon_{1t}, \epsilon_{2t}) \neq 0$.In Q1 and Q2 we also assume that both W_t and Z_t are $I(0)$ variables, and that the reduced form of (6) and (7) is a stationary VAR.

1. Discuss the identification of the two structural equations.

   Answer:

   Order and rank conditions. The two equations are exactly identified on the order condition. Under the assumption that both W_t and Z_t are relevant variables, the rank conditions is satisfied (the last sentence is more or less obvious given the question, so many will drop it).

2. Assuming that the first equation in the SEM is identified:

   (a) Explain why OLS gives inconsistent estimation of β_{10} (and the other coefficients).

   Answer: Simultaneity bias as X_t is in general correlated with ϵ_{1t}. For exampel if ϵ_{1t} is increased, not only Y_t is increased, but X_t also changes (assuming as $\gamma_{10} \neq 0$)

   (b) Explain why IV estimation gives consistent estimators.

   Answer: IV is the same as replacing X_t in (6) by the OLS predicted variables $\hat{X}_t$, based on the reduced form, and estimating by OLS. The resulting estimator is called 2SLS. The correlation is "broken" as $\hat{X}_t$ is uncorrelated with ϵ_t. 2SLS works in the same way in the case of overidentification.

We now change the assumption about stationary of W_t and Z_t, and assume instead that both Z_t and W_t are $I(1)$ variables.

3. Explain why Y_t and X_t are in general $I(1)$ variables when Z_t and W_t are $I(1)$.

   Answer: This follows from the reduced form, and from the assumption that W_t and Z_t (or at least one of them) has non-zero coefficients in the reduced form. Logically, Y_t and X_t are then also $I(1)$.

4. We next consider system under the assumptions: $\beta_{10} = \gamma_{10} = \gamma_{11} = \gamma_{20} = 0$ and $\alpha = 1$ (but $Cov(\epsilon_{1t}, \epsilon_{2t}) \neq 0$ as before). Explain how you use the results in Table 1 to test the null hypothesis of no cointegration.

   Answer: With reference to the critical values in the Ericsson and MacKinnon paper, we conclude that the H_0 is rejected at the 5 % frequency. The battery of misspecification test gives reason to believe the that assumptions of the test are reasonabley realistic for the DGP. Based in the information given, X_t is weakly exogenous with respect to the cointegration parameters.

   It is also possible to add: Under the null of no-cointgation, negative t-values of the estimated ECM coefficients are much more probable than in case with stationary variables. This means that the use of critical values from the normal or t-distribution is hazardous and is the source of the spurious regression problem. By using DF-type disgributions that account for the number of $I(1)$ variables in the model, we are able to make formally correct inference, holding the Type-1 error probabilities close to the chosen significance level.

5. Finally, consider the model (6) and (7) under the assumptions: $\beta_{20} = \gamma_{20} = 0$. Explain briefly what properties the reduced form must have for the two endogenous variables to be $I(1)$ in this case. How many cointegration relationships can we have, at most, in this case?

   Answer: The RF must have at least one unit root, but it can also have two. We can have at most one ci-relationship, in which case there is only one unit-root in the reduced form.

```
     Modelling DY by OLS

     The estimation sample is: 3 - 101

                    Coefficient  Std.Error  t-value  t-prob
Constant              -0.123867    0.08941    -1.39  0.1692
DX                     0.252379    0.08154     3.10  0.0026
Y_1                   -0.558797    0.08878    -6.29  0.0000
X_1                    0.388778    0.06430     6.05  0.0000
Z                      0.799442    0.07869     10.2  0.0000

sigma                  0.801969  RSS                60.4565135
R^2                    0.683199  F(4,94) =     50.68 [0.000]**
Adj.R^2                0.669718  log-likelihood       -116.062
no. of observations          99  no. of parameters           5
mean(DY)              -0.319047  se(DY)                1.39545

AR 1-2 test:      F(2,92)    =   1.7649 [0.1770]
ARCH 1-1 test:    F(1,97)    =  0.39606 [0.5306]
Normality test:   Chi^2(2)   =  0.43364 [0.8051]
Hetero test:      F(8,90)    =  0.73809 [0.6576]
Hetero-X test:    F(14,84)   =  0.69692 [0.7713]
```

Table 1: ECM type model for Y_t ..Difference is symbolized by D, i.e, $DY = \Delta Y_t$ and $DX = \Delta X_t$