UNIVERSITY OF OSLO
DEPARTMENT OF ECONOMICS

Postponed exam: **ECON4160 – Econometrics – Modeling and Systems Estimation**

Date of exam: Thursday, January 11, 2018

Time for exam: 09:00 a.m. – 12:00 noon

The problem set covers 7 pages (incl. cover sheet)

Resources allowed:

- Open book exam, where all written and printed resources – as well as calculator - is allowed

The grades given: A-F, with A as the best and E as the weakest passing grade. F is fail.

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

Day of exam: 11 January 2018

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 %)

Assume that you have a sample of n mutually independent triplets (X_i, Y_i, Z_i) that are identically distributed. Denote the empirical (sample) variances and covariances by $\hat{\sigma}_X^2$, $\hat{\sigma}_Y^2$, $\hat{\sigma}_Z^2$, $\hat{\sigma}_{XZ}$, $\hat{\sigma}_{XY}$, $\hat{\sigma}_{YZ}$.

Assume that the task is to estimate the relationship between Y and X, and that the relationship is specified to be linear in both parameters and variables.

1. What is the expression for the OLS estimator of the slope coefficient, denoted by γ, in the relationship between Y and X?

2. What is the expression for the IV estimator of γ, when Z is the instrumental variable?

3. Under which assumptions is the IV estimator a consistent estimator of γ?

4. The ratio between $Var(\hat{\gamma}_{OLS})$ and $Var(\hat{\gamma}_{IV})$ can be expressed as:

   $$\frac{Var(\hat{\gamma}_{OLS})}{Var(\hat{\gamma}_{IV})} = \frac{(\hat{\sigma}_{XZ})^2}{\hat{\sigma}_X^2 \hat{\sigma}_Z^2} \tag{1}$$

   for the case where X_i and Z_i are deterministic variables. (You do not need to show (1)).

   What does the expression in (1) tell us about the statistical efficiency of the IV estimation procedure?

Question B (40 %)

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

$$E(\varepsilon_t) = 0 \tag{2}$$

$$Var(\varepsilon_t) = \sigma_\varepsilon^2 > 0 \tag{3}$$

$$Cov(\varepsilon_t, \varepsilon_{t-j}) = 0, \text{ for } j = \pm1, \pm2, ... \tag{4}$$

Assume that the DGP for the time series Y_t is:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t. \tag{5}$$

Based on (2)-(5), answer the following questions:

1. What are the expressions for $E(Y_t \mid Y_{t-1}, Y_{t-2})$ and $Var(Y_t \mid Y_{t-1}, Y_{t-2})$?
2. Give the expression for $E(Y_t)$ in the case where Y_t is a covariance stationary time series, denoted $Y_t \sim I(0)$.
3. Assume that $\phi_1 + \phi_2 = 1$. Explain why Y_t is a unit-root non stationary time series, $Y_t \sim I(1)$, in this case.
4. Give an expression for the DGP equation when $\phi_1 + \phi_2 = 1$. What is the condition for stationarity of ΔY_t in this case?
5. Figure 1 (at the back of the question set) shows time plots of two time series named Y_{1t} andY_{2t}, and Table 1 and Table 2 show PcGive output for unit-root tests for the two time series. Use the provided information to decide which (if any) of the series can be characterized as $I(0)$.

Question C (40 %)

Consider the following model equation for Norwegian inflation, INF_t:

$$\begin{aligned} INF_t = {} & \phi_0 + \phi_1 INF_{t-1} + \beta_1 U_t + \beta_2 U_{t-1} \\ & + \beta_3 IMP_t + \beta_4 D_{INFt} + \epsilon_t, \end{aligned} \tag{6}$$

where the right hand side variables are the rate of unemployment, U_t, imported inflation, IMP_t, and a dummy variable, D_{INFt} (capturing the effects of breaks in the relationship between 1904 and 1948). ϵ_t is the error term of the model equation.

We have annual data that allows us to estimate the relationship from 1904 to 2015. The three variables INF_t, $IINF_t$ and U_t are measured in percentage points.

1. Show that (6) can be re-parameterized as:

$$\begin{aligned}(7) \qquad \Delta INF_t = {} & \phi_0 + \pi(INF_{t-1} - IMP_t) + \beta_1 \Delta U_t + \gamma_U U_{t-1} \\ & + \gamma_{IMP} IMP_t + \beta_4 D_{INFt} + \epsilon_t,\end{aligned}$$

   where Δ is the difference operator, and:

$$\begin{aligned}\pi &= \phi_1 - 1, \\ \gamma_U &= \beta_1 + \beta_2, \\ \gamma_{IMP} &= \beta_3 + \pi.\end{aligned}$$

   The estimated version of (7) is:

$$\begin{aligned}\Delta INF_t = {} & -\underset{(0.229)}{0.6652}\ \Delta U_t - \underset{(0.0326)}{0.597}\ (INF_{t-1} - IMP_t) \\ (8) \qquad & -\underset{(0.0863)}{0.238}\ U_{t-1} - \underset{(0.0331)}{0.2474}\ IMP_t \\ & + \underset{(0.416)}{1.862} + \underset{(0.0552)}{1.002}\ D_{INFt}\end{aligned}$$

   Method: OLS, Sample: 1904-2015, (112 observations)
   $\hat{\sigma} = 1.94804$, $RSS = 402.25405$

   where the numbers below the point estimates of the coefficients are estimated standard errors. We have also included information about estimation method, OLS, the estimation sample used, the estimated standard deviation of the error term ($\hat{\sigma}$), and the residual sum of squares (RSS).

2. To save space we have not included any mis-specification tests in (8), but they do not indicate any mis-specification. The inclusion of D_{INFt} is important for this result. Why?

3. Show that (7) implies a long-run relationship which we can write as:

$$INF = \alpha_0 + \alpha_1 IMP + \alpha_2 U \tag{9}$$

   and use the results in (8) to obtain point estimates of the coefficients in (9).

4. In order to be able to solve (9) for an equilibrium rate of unemployment which is independent of the level of inflation (often called natural rate of unemployment), we would need α_1 to be equal to 1. Test this restriction statistically by using the results in (8).

5. A critique of (8) is that in particular the estimate for β_1 is affected by simultaneity bias. Explain why an answer to this critique is to estimate (7) by the IV procedure, instead of by OLS.

6. Equation (10) shows IV estimation results when a variable that captures breaks in unemployment (between 1904 and 1948) is used as an instrumental variable (This variable is practically speaking uncorrelated with D_{INFt}). Compare the results with the OLS based results above and give your comments.

$$\begin{aligned}
\Delta INF_t = \; & - \underset{(0.319)}{1.236} \; \Delta U_t - \underset{(0.0345)}{0.5759} \; (INF_{t-1} - IMP_t) \\
& - \underset{(0.0899)}{0.2744} \; U_{t-1} - \underset{(0.034)}{0.2483} \; IMP \\
& + \underset{(0.431)}{1.984} + \underset{(0.0568)}{0.998} \; D_{INFt}
\end{aligned} \tag{10}$$

Method: IV, Sample: $1904 - 2015$, (112 observations)
$\hat{\sigma} = 2.00439$, $RSS = 425.862846$

Graphs and tables to Question B

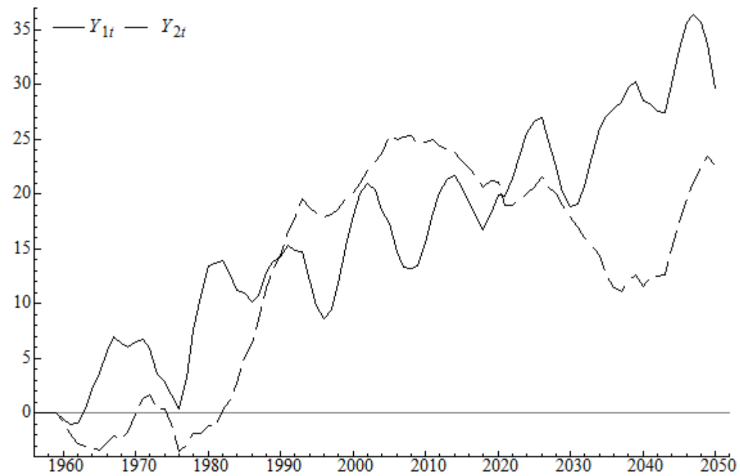

Figure 1: Time plots of Y_{1t} and Y_{2t} (dashed line).

```
The sample is: 1960 - 2050 (94 observations and 2 variables)

Y1: ADF tests (T=91, Constant; 5%=-2.89 1%=-3.50)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 2     -1.820       0.97710    1.085     -3.671  0.0004
 1     -2.693       0.96505    1.160      9.901  0.0000
 0     -1.168       0.97821    1.676

Y2: ADF tests (T=91, Constant; 5%=-2.89 1%=-3.50)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob
 2     -1.445       0.98624   0.8856      1.121  0.2653
 1     -1.342       0.98726   0.8869      7.112  0.0000
 0    -0.8967       0.98938    1.107
```

Table 1: Augmented Dickey Fuller tests for Y_{1t} and Y_{2t} (shown in Figure 1). The Dickey-Fuller regression includes a constant.

The sample is: 1960 - 2050 (94 observations and 2 variables)

Y1: ADF tests (T=91, Constant+Trend; 5%=-3.46 1%=-4.06)

D-lag	t-adf	beta Y_1	sigma	t-DY_lag	t-prob
2	-6.749**	0.70752	0.8980	0.7284	0.4684
1	-8.481**	0.72875	0.8956	14.36	0.0000
0	-2.633	0.85196	1.634		

Y2: ADF tests (T=91, Constant+Trend; 5%=-3.46 1%=-4.06)

D-lag	t-adf	beta Y_1	sigma	t-DY_lag	t-prob
2	-1.644	0.97766	0.8867	1.230	0.2222
1	-1.459	0.98039	0.8893	7.128	0.0000
0	-3.7159	0.98799	1.113		

Table 2: Augmented Dickey Fuller tests for Y_{1t} and Y_{2t} (shown in Figure 1).The Dickey-Fuller regression includes both constant and trend.