UNIVERSITY OF OSLO
DEPARTMENT OF ECONOMICS

Postponed exam: **ECON4160 – Econometric Modelling and System Estimation**

Date of exam: Friday, December 17, 2010 **Grades are given:** January 6, 2011

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

The problem set covers 4 pages (incl. cover sheet)

Resources allowed:

- All written and printed resources, as well as calculator, are 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: Econometric Modelling and System Estimation

Day of exam: 17 December, 2010, postponed exam

Time of day: 9:00-12:00

This is a 3 hour school exam.

Guidelines:
In the grading, each of the 3 questions will count 1/3 .

1. Consider the following structural model:

$$\begin{aligned} &(1) \qquad Y_{1i} = \beta_0 + \beta_1 Y_{2i} + \beta_2 X_i + \varepsilon_i, \quad 0 \leq \beta_1 < 1 \\ &(2) \qquad Y_{2i} = Y_{1i} + X_i \\ &\qquad E(\varepsilon_i) = 0, \ Var(\varepsilon_i) = \sigma_\varepsilon^2 > 0. \end{aligned}$$

where $i = 1, \ldots, n$, and the endogenous variables are Y_{1i} and Y_{2i}. The assumptions about the properties of the disturbances ε_i are conditional on X_i. For concreteness we can define Y_1 as private consumption, Y_2 as GDP, and X as autonomous expenditure. (1) is therefore interpretable as a consumption function and (2) can be interpreted as a simplified equilibrium condition for the aggregate product market of a closed economy.

(a) Explain how you can test the hypothesis that there is no relationship between X and Y_1. Base your answer on a complete specification of the econometric model.

In the rest of question 1 we assume that $0 < \beta_1 < 1$ and $\beta_2 = 0$.

(b) Set $E(X_i) = \mu_X$ and $Var(X_i) = \sigma_X^2 > 0$. Use the reduced form of the structural model (1) and (2) to show that

$$\begin{aligned} \mu_{Y_1} &= \tfrac{\beta_0+\beta_1\mu_X}{1-\beta_1} & \mu_{Y_2} &= \tfrac{\beta_0+\mu_X}{1-\beta_1} \\ \sigma_{Y_1,Y_2} &= \tfrac{\beta_1\sigma_X^2+\sigma_\varepsilon^2}{(1-\beta_1)^2} & \sigma_{Y_2}^2 &= \tfrac{\sigma_X^2+\sigma_\varepsilon^2}{(1-\beta_1)^2} \\ \sigma_{Y_2,X} &= \tfrac{\sigma_X^2}{1-\beta_1} & \sigma_{Y_2\varepsilon} &= \tfrac{\sigma_\varepsilon^2}{1-\beta_1} \end{aligned}$$

where μ_{Y_1} and μ_{Y_2} are expectations, σ_{Y_1,Y_2}, $\sigma_{Y_2,X}$ and $\sigma_{Y_2,\varepsilon}$ are covariances, and $\sigma_{Y_2}^2$ is the variance of Y_2.

(c) Assume that we want to estimate β_1. Show that the asymptotic bias of the OLS estimator $\hat{\beta}_{1,OLS}$ based on (1) with $\beta_2 = 0$, is:

$$\text{plim}(\hat{\beta}_{1,OLS}) - \beta_1 = (1 - \beta_1)\,\theta > 0,$$

where

$$\theta \equiv \frac{\sigma_\varepsilon^2}{\sigma_X^2 + \sigma_\varepsilon^2}.$$

Try also to explain this result intuitively.

(d) Can anything definite be said about the finite sample bias of $\hat{\beta}_{1,OLS}$? Explain briefly.

(e) Show that the instrumental variable estimator

$$\hat{\beta}_{1,IV} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(Y_{1i} - \bar{Y}_1\right)\left(X_i - \bar{X}\right)}{\frac{1}{n}\sum_{i=1}^{n}\left(Y_{2i} - \bar{Y}_2\right)\left(X_i - \bar{X}\right)}$$

is a consistent estimator of β_1.

(f) Show that $\hat{\beta}_{1,2SLS} = \hat{\beta}_{1,IV}$, where $\hat{\beta}_{1,2SLS}$ is the two stage least squares estimator.

2. Consider the following model of joint equilibrium in two markets:

$$Q_{1i} = \beta_{10} + \beta_{11}P_{1i} + \beta_{12}P_{2i} + \beta_{13}X_i + \varepsilon_{1i} \tag{3}$$
$$Q_{1i} = \beta_{20} + \beta_{21}P_{1i} + \beta_{24}W_i + \varepsilon_{2i} \tag{4}$$
$$Q_{2i} = \beta_{30} + \beta_{31}P_{1i} + \beta_{32}P_{2i} + \beta_{33}X_i + \varepsilon_{3i} \tag{5}$$
$$Q_{2i} = \beta_{40} + \beta_{42}P_{2i} + \varepsilon_{4i} \tag{6}$$

(3) and (4) are the demand and supply equations for "market 1". The demand equation of "market 2" is (5) and the supply equation for this market is (6). The variables Q_j and P_j $(j = 1, 2)$ are equilibrium quantities (Q_j) and equilibrium prices (P_j) in the two markets. X_i and W_i are exogenous variables. ε_{ji} $(j = 1, 2, 3, 4)$ are disturbances.

Discuss the identification properties of the model.

3. We have a data set with information about Norwegian women's years of education. The data set also contains variables that may be correlated with education length. The sample size is 50487 individuals. A brief description of the variables in the data set is:

- $LEDUC_i$: The natural logarithm of the number of years of education for individual i.
- $CHILD_i$: The number of children (0 to 18 years old) for individual i.
- $LFINC_i$: The natural logarithm of annual labour-free income (at constant 2005 prices) for individual i.
- JOB_i: Indicator variable which takes the value 1 if the woman i has a job, and 0 if she is unemployed. All the women in the sample are in the labour force.
- $NOURBAN_i$: Indicator variable which takes the value 1 if the woman i lives in a non-urban area.

Consider the regression output (from Pc-Give) on the next page:

```
EQ( 1) Modelling LEDUC by OLS-CS

       The estimation sample is: 1 - 50487

                  Coefficient  Std.Error  t-value   t-HCSE
Constant              3.80344    0.06513     58.4   47.356
NONURBAN          -0.0360801   0.002232    -16.2  -16.575
JOB                0.0918585   0.006110     15.0   17.385
LFINC              -0.348815    0.01201    -29.0  -23.763
LFINC*LFINC        0.0181051  0.0005651     32.0   26.632
CHILD              0.0655028    0.01491     4.39   4.1005
CHILD*CHILD       -0.0114857  0.0006749    -17.0  -16.793
CHILD*LFINC      0.000440782   0.001222    0.361  0.33652

sigma                0.208972  RSS                 2204.39039
R^2                  0.101614  F(7,50479) =  815.6 [0.000]**
Adj.R^2              0.101489  log-likelihood         7406.14
no. of observations     50487  no. of parameters            8


Normality test:   Chi^2(2)   =   4323.7 [0.0000]**
Hetero test:      F(11,50475)=   31.846 [0.0000]**
Hetero-X test:    F(22,50464)=   21.665 [0.0000]**
```

The $t - HSCE$ column contains t-values that are based on heteroscedastic consistent estimators of the coefficient standard errors.

(a) Give your assessment of the importance of the different explanatory variables.

(b) The mean of labour-free income is 15000 kroner and the average length of education in the sample is 12 years. What is the estimated derivative of education length with respect to labour free income, when these sample means are used?

(c) For which values of $CHILD$ is the estimated derivative of education length with respect to the number of children positive, and for which values of $CHILD$ is the derivative negative?

(d) Assume that you are asked to investigate a different hypothesis, namely that the probability of having a job is a function of the variable $CHILD$. Explain briefly how you would test this hypothesis with the aid of an econometric model.