# *UNIVERSITY OF OSLO*
# *DEPARTMENT OF ECONOMICS*

Postponed exam: **ECON4160 – Econometrics – Modeling and systems estimation**

Date of exam: Friday, August 14, 2009

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

The problem set covers 5 pages

Resources allowed:

- 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: Econometric Modelling and System Estimation

**Day of exam:** 14 August, 2009. Postponed exam.

**Time of day:** 9:00-12:00

This is a 3 hour school exam.

**Guidelines**:
In the grading, each of the 4 questions will count 25 %.

1. (a) Show that the equation

(1) $$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i,\ i = 1, 2, 3, ...., n,$$

can be written as

(2) $$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{y}, \boldsymbol{\beta}$ and $\boldsymbol{\varepsilon}$ denote vectors, and $\mathbf{X}$ is a suitably defined matrix.

(b) Explain why the OLS estimator $\widehat{\boldsymbol{\beta}}$ of the parameter vector $\boldsymbol{\beta}$ satisfies the "normal equations", which in matrix notation are given by:

(3) $$(\mathbf{X}'\mathbf{X})\widehat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}.$$

(c) Consider re-writing (1) as

(4) $$y_i = \alpha + \beta_2(x_{2i} - \bar{x}_2) + \beta_3(x_{3i} - \bar{x}_3) + \varepsilon_i,\quad i = 1, 2, 3, ...., n,$$

with $\alpha = \beta_1 + \beta_2\bar{x}_2 + \beta_3\bar{x}_3$, where $\bar{x}_2$ and $\bar{x}_3$ denote the means of $x_{2i}$ and $x_{3i}$ ($\bar{x}_k = (1/n)\sum_{i=1}^n x_{ki},\ k = 2, 3$). Show that the "normal equations" corresponding to equation (4) can be expressed as:

$$\begin{pmatrix} n & 0 & 0 \\ 0 & \sum_{i=1}^n (x_{2i} - \bar{x}_2)^2 & \sum_{i=1}^n (x_{2i} - \bar{x}_2)(x_{3i} - \bar{x}_2) \\ 0 & \sum_{i=1}^n (x_{2i} - \bar{x}_2)(x_{3i} - \bar{x}_2) & \sum_{i=1}^n (x_{3i} - \bar{x}_3)^2 \end{pmatrix} \begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta}_2 \\ \widehat{\beta}_3 \end{pmatrix}$$
$$= \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n (x_{2i} - \bar{x}_2) y_i \\ \sum_{i=1}^n (x_{3i} - \bar{x}_3) y_i \end{pmatrix}.$$

(d) Show that the OLS estimates $\widehat{\beta}_2$ and $\widehat{\beta}_3$ are given by the two equations:

(5) $$s_{x_2}^2 \widehat{\beta}_2 + s_{x_2 x_3}\widehat{\beta}_3 = s_{x_2 y}, \text{ and}$$

(6) $$s_{x_2 x_3}\widehat{\beta}_2 + s_{x3}^2 \widehat{\beta}_3 = s_{x_3 y}$$

where

$$s_{x_k}^2 = \frac{1}{n}\sum_{i=1}^n (x_{ki} - \bar{x}_k)^2,\ k = 2, 3$$

$$s_{x_2 x_3} = \frac{1}{n}\sum_{i=1}^n (x_{2i} - \bar{x}_2)(x_{3i} - \bar{x}_3)$$

$$s_{x_k y} = \frac{1}{n}\sum_{i=1}^n (y_i - \bar{y})(x_{ki} - \bar{x}_k),\ k = 2, 3.$$

(e) Explain why the two parameters $\beta_2$ and $\beta_3$ can only be estimated if the degree of correlation between $x_2$ and $x_3$ is less than perfect in the sample.

2. Assume that the disturbances $\varepsilon_i$ in (1) have the so-called *classical properties*, in particular that the variance $\sigma^2$ is the same for all disturbances, and that the explanatory variables are stochastic.

   (a) Show that the conditional expectation of the OLS estimator $\widehat{\boldsymbol{\beta}}$ defined in (3) is $\boldsymbol{\beta}$, and that the conditional variance of the same estimator is $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

   (b) Explain briefly why the unconditional expectation of $\widehat{\boldsymbol{\beta}}$ is $\boldsymbol{\beta}$.

   (c) Show that the variances for the OLS estimators for the parameters $\beta_2$ and $\beta_3$ are given by:

$$Var[\widehat{\beta}_k] = \frac{\sigma^2}{ns_{x_k}^2(1 - r_{x_2x_3}^2)}, \; k = 2, 3 \qquad (7)$$

   where $r_{x_2x_3}^2$ is the squared correlation coefficient between $x_2$ and $x_3$.

   (d) Show that the OLS estimator of the parameter vector $\widehat{\boldsymbol{\beta}}$ in (3) is consistent.

3. Below we give four different structures of a partial market equilibrium model:

$$\begin{aligned}
& \textit{Structure 1} \\
Q_t &= 15 - P_t + \varepsilon_{dt} \\
Q_t &= -0.2 + 0.5P_t + \varepsilon_{st} \\
& \textit{Structure 2} \\
Q_t &= 15 - P_t + 1.5X_{dt} + \varepsilon_{dt} \\
Q_t &= -0.2 + 0.5P_t + \varepsilon_{st} \\
& \textit{Structure 3} \\
Q_t &= 15 - P_t + \varepsilon_{d,t} \\
Q_t &= -0.2 + 0.5P_t + 2X_{st} + \varepsilon_{st} \\
& \textit{Structure 4} \\
Q_t &= 15 - P_t + 1.5X_{dt} + \varepsilon_{dt} \\
Q_t &= -0.2 + 0.5P_t + 2X_{st} + \varepsilon_{st}
\end{aligned}$$

   The endogenous variable, $P_t$, is the equilibrium price, and the other endogenous variable, $Q_t$, is equilibrium quantity. $X_{dt}$ and $X_{st}$ are variables that are exogenous in the econometric sense. The disturbances $\varepsilon_{dt}$ $(t = 1, 2, ..., T)$ are independent normally distributed with zero means, and the same is the case for $\varepsilon_{st}$, The disturbances of the two equations are uncorrelated, i.e. $Cov(\varepsilon_{st}, \varepsilon_{dt}) = 0$ for all $t$.

   (a) Discuss the identification of these 4 structural models.

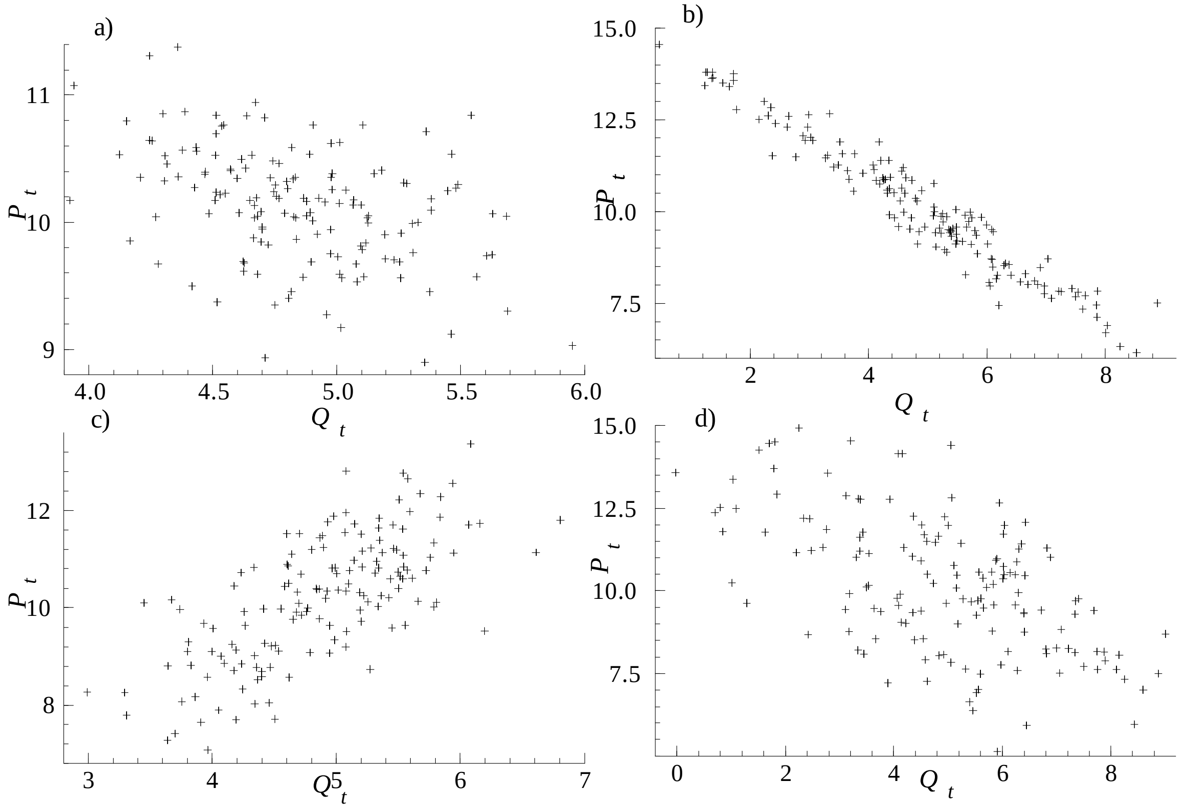


Figure 1: Scatter plots of simulated data corresponding to to Structure 1-4 in Question 3.

(b) Figure 1 shows four scatter plots of artificial (simulated) data of price ($P_t$) and quantity ($Q_t$). Make a guess about which graph ( a), b) c) and d)) "belong to" which structure (1, 2, 3 or 4). Explain briefly your reasoning.

(c) Explain how you would estimate the equations that are identified in the different structures.

4. Consider a different structure from the ones we consider in question 3, namely

$$Q_t = 15 - P_t + 1.5X_{dt} + \varepsilon_{dt}, \quad (8)$$
$$Q_t = -0.2 + 0.5P_{t-1} + \varepsilon_{st}, \quad (9)$$

but where we make the same assumptions about $X_{dt}$, $\varepsilon_{dt}$ and $\varepsilon_{st}$ as we do in question 3.

(a) Explain why OLS estimation of a two-equation model, where first $Q_t$ is regressed on a *Constant* and $P_{t-1}$, and, second, $P_t$ is regressed on a *Constant*, and $Q_t$ and $X_{dt}$, give consistent estimates of the parameters in (8) and (9).

(b) Are the OLS estimates also unbiased for a finite sample size? Explain.

(c) Consider the result from one of the regressions mentioned in question 4(a):

$$Q_t = \underset{(0.2677)}{-0.194} + \underset{(0.0259)}{0.4979}\ P_{t-1} \tag{10}$$

OLS, (Sample is 1-150)

and also the result from IV estimation

$$Q_t = \underset{(0.3261)}{-0.080} + \underset{(0.0317)}{0.4867}\ P_{t-1} \tag{11}$$

$IV$, (Sample is 1-150)

where $X_{dt-1}$ is used as the instrument for $P_{t-1}$. The correlation coefficient between $X_{dt-1}$ and $P_{t-1}$ is 0.81. The numbers in parentheses are estimated standard errors.

Are the differences between the two estimation results as you would expect, and which estimation method would you prefer to use? Explain your reasoning.

(d) Consider again (8) and (9), and imagine that you as an econometrician know the model specification, but you do not know the coefficient values. Hence the model that you specify is

$$Q_t = \gamma_{do} + \gamma_{dp} P_t + \gamma_{dx} X_{dt} + \varepsilon_{dt}, \tag{12}$$

$$Q_t = \gamma_{so} + \gamma_{sp} P_{t-1} + \varepsilon_{st}, \tag{13}$$

where the two disturbances have the same properties as given in connection with Question 3 above (normally distributed and uncorrelated). Explain how you would solve the following tasks:

i. Test the hypothesis that a positive change in $X_{dt}$ that lasts for one period has a positive effect on the market price in the same period as the shock occurs.

ii. Test the hypothesis that a *permanent* change in $X_{dt}$ has a long-run effect on the market price that is different from the short-run effect of this change.

iii. The ratio $(-\gamma_{sp}/\gamma_{dp})$ is an important parameter of interest in this model. Sketch how you can estimate a confidence interval for this parameter.