

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

Postponed exam: ECON4160 – Econometrics – Modeling and systems estimation

Date of exam: Tuesday, August 08 2006

Time for exam: 9:00 a.m. – 12:00 noon

The problem set covers 5 pages (included cover page)

Resources allowed:

- All written and printed resources, as well as calculator

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

ECON 4160: ECONOMETRICS –  
MODELLING AND SYSTEMS ESTIMATION  
POSTPONED EXAM, SPRING 2006

PROBLEM 1 (WEIGHT: 25 %)

Consider the following simple macro model:

$$\begin{aligned}C_t &= a_0 + a_1 Y_t + u_t, \\I_t &= b_0 + b_1 Y_t + b_2 Y_{t-1} + v_t, \\Y_t &= C_t + I_t + X_t,\end{aligned}$$

where

- $Y_t$  = Gross national product in year  $t$ ,
- $C_t$  = Private consumption expenditure in year  $t$ ,
- $I_t$  = Gross investment expenditure in year  $t$ ,
- $X_t$  = Public expenditure + Exports – Imports in year  $t$ ,
- $u_t, v_t$  = Disturbances.

We assume that  $X_t$  is exogenous and that  $Y_{t-1}$  is predetermined (which for the present purpose means that it has a similar econometric status as  $X_t$ ). Answer briefly the following questions and state in each case the reason for your answer:

- (1a)** Specify the reduced form of this macro model and determine which of its equations are exactly identified and which are overidentified.
- (1b)** Which method(s) would you use to estimate the model's equations? Explain briefly.
- (1c)** Would your answers to **(1a)** and **(1b)** have been different if your data set had contained time series for the three components of  $X_t$ , i.e., Public expenditure, Exports, and Imports separately, rather than containing time series for the aggregate  $X_t$  only?

[PROBLEM SET CONTINUES ON NEXT PAGE]

**PROBLEM 2** (WEIGHT: 25 %)

Consider the following two-equation time series model for explaining the simultaneous determination of wage inflation ( $w_t$ ) and consumer price inflation ( $p_t$ ) in the United Kingdom ( $t$  denotes year):

$$w_t = \alpha_1 + \beta_1 p_t + \gamma_1 p_{t-1} + \gamma_2 q_t + \varepsilon_t, \quad (\text{wage equation})$$

$$p_t = \alpha_2 + \beta_2 w_t + \gamma_3 x_t + \gamma_4 m_t + \gamma_5 m_{t-1} + \delta_t, \quad (\text{price equation})$$

where

$w_t$  = Rate of increase of wage index,

$p_t$  = Rate of increase of consumer price index,

$m_t$  = Rate of increase of import price index,

$x_t$  = Rate of increase of labour productivity,

$q_t$  = 100\*Number of unfilled jobs/Number of employees,

$\varepsilon_t, \delta_t$  = Disturbances,

The model has been estimated from data for the years 1951–1969 by using, for both equations, the Ordinary Least Squares (OLS) and the Two-Stage Least Squares (2SLS). The point estimates are (standard error estimates are suppressed):

Estimate of	OLS	2SLS
Wage eq.:		
$\alpha_1$	0.276	0.272
$\beta_1$	0.258	0.257
$\gamma_1$	0.046	0.046
$\gamma_2$	4.959	4.966
Price eq.:		
$\alpha_2$	2.693	2.686
$\beta_2$	0.232	0.233
$\gamma_3$	-0.544	-0.544
$\gamma_4$	0.247	0.246
$\gamma_5$	0.064	0.046
$R^2$ of Wage eq.	0.924	0.920
$R^2$ of Price eq.	0.982	0.981

Discuss critically the following statements:

**(2a)** “Since the equations estimated by OLS have highest  $R^2$ , the OLS estimates should be preferred”.

**(2b)** “Since the OLS and the 2SLS results are practically identical, using 2SLS is not meaningful”.

[PROBLEM SET CONTINUES ON NEXT PAGE]

### PROBLEM 3 (WEIGHT: 50 %)

In analyzing econometrically relationships between consumption and income from data from individual households, it is often a problem that these variables are imperfectly measured. We may, however, in addition to error-ridden measures of consumption and income, have observations on other variables which can be assumed to be correlated with the true values of these improperly measured variables. We therefore consider the following measurement error model,  $i$  indexing household:

$$\begin{aligned}(1) \quad & \eta_i = \alpha + \beta\xi_i, \\(2) \quad & y_i = \eta_i + \varepsilon_i, \\(3) \quad & x_i = \xi_i + \delta_i, \\(4) \quad & \xi_i = \lambda_x + \gamma_x q_i + u_i, \quad i = 1, \dots, n,\end{aligned}$$

where  $(\eta_i, \xi_i)$  denote true, unobserved consumption and income, respectively,  $(y_i, x_i)$  are their observed counterparts,  $(\varepsilon_i, \delta_i)$  are measurement errors,  $q_i$  is an exogenous variable which determines true consumption, say age, wealth, or education, and  $u_i$  is a disturbance. We assume [IID( $\mu, \theta^2$ )] means identically, independently distributed with expectation  $\mu$  and variance  $\theta^2$

$$(5) \quad \begin{aligned} \varepsilon_i | q_i &\sim \text{IID}(0, \sigma_\varepsilon^2), & \delta_i | q_i &\sim \text{IID}(0, \sigma_\delta^2), & u_i | q_i &\sim \text{IID}(0, \sigma_u^2), \\ \varepsilon_i, \delta_i, u_i &\text{ are uncorrelated,} & & & i &= 1, \dots, n. \end{aligned}$$

**(3a)** The model contains four equations: Which are its four endogenous variables, and which of them are observable and which are latent? The consumption function (1) is specified without a disturbance. Have you any comment to this simplifying assumption?

**(3b)** Express  $\text{var}(y_i)$ ,  $\text{var}(x_i)$ ,  $\text{cov}(y_i, x_i)$ ,  $\text{cov}(y_i, q_i)$ , and  $\text{cov}(x_i, q_i)$  by means of the parameters in (1)–(5) and  $\sigma_q^2 = \text{var}(q_i)$ .

**(3c)** Derive from (1)–(3), by eliminating  $\eta_i$  and  $\xi_i$ , an equation between  $y_i$  and  $x_i$ . Explain why  $q_i$  satisfies the requirements for being a valid instrumental variable for either of  $y_i$  and  $x_i$  this equation.

For estimating the marginal propensity to consume of latent income,  $\beta$ , it has been proposed to use an estimator of the following form:

$$\widehat{\beta}(z) = \frac{M[y, z]}{M[x, z]},$$

where  $M[y, z]$  and  $M[x, z]$  denote, respectively, the empirical covariance between  $y$  and  $z$  and between  $x$  and  $z$ , and  $z_i$  is a so far unspecified, but observable, variable.

[PROBLEM SET CONTINUES ON NEXT PAGE]

**(3d)** Derive the probability limit of  $\widehat{\beta}(z)$  for the following four choices of  $z_i$ :

(i)  $z_i = x_i$ .

(ii)  $z_i = y_i$ .

(iii)  $z_i = \widehat{x}_i = \widehat{\lambda}_x + \widehat{\gamma}_x q_i$ , where  $(\widehat{\lambda}_x, \widehat{\gamma}_x)$  are obtained by OLS regression of  $x_i$  on  $q_i$ .

(iv)  $z_i = \widehat{y}_i = \widehat{\lambda}_y + \widehat{\gamma}_y q_i$ , where  $(\widehat{\lambda}_y, \widehat{\gamma}_y)$  are obtained by OLS regression of  $y_i$  on  $q_i$ .

How would you characterize the estimators obtained in cases (i)–(iv) and their properties? Why are  $(\widehat{\lambda}_x, \widehat{\gamma}_x)$  unbiased estimators of  $(\lambda_x, \gamma_x)$ ?

Finally, assume that equations (1)–(3) are part of a simultaneous household model containing a total of  $K$  observable, exogenous variables whose values for household  $i$  are:  $q_{1i}, \dots, q_{Ki}$ . These  $K$  variables may include age, net wealth, the number of household members, their education, etc. The reduced-form equations for true consumption  $\eta_i$  and true income  $\xi_i$  therefore have the form:

$$\begin{aligned}\eta_i &= \Pi_{y0} + \sum_{j=1}^K \Pi_{yj} q_{ji} + v_i, \\ \xi_i &= \Pi_{x0} + \sum_{j=1}^K \Pi_{xj} q_{ji} + u_i,\end{aligned}$$

where the  $\Pi$ 's denote reduced-form coefficients and  $u_i$  and  $v_i$  are disturbances. Let their OLS estimators be denoted by a tilde ( $\widetilde{\phantom{x}}$ ).

**(3e)** It has been proposed to estimate  $\beta$  by using the estimator  $\widehat{\beta}(z)$  with

$$(v) \quad z_i = \widetilde{\Pi}_{y0} + \sum_{j=1}^K \widetilde{\Pi}_{yj} q_{ji}$$

$$(vi) \quad z_i = \widetilde{\Pi}_{x0} + \sum_{j=1}^K \widetilde{\Pi}_{xj} q_{ji}$$

Comment on these proposed choices of  $z_i$ . Would you prefer them to (iii) and (iv)?