Exam Advanced Econometrics 2 (FEB-UvA: 2005/2006)
resit of 28 March, 2006; 14.00-16.00 hrs.

This is NOT an "open-book" exam; apart from pen and paper no further aids and tools are allowed to be used. Write your name and student number on all sheets that you hand in for marking. For each separate (sub-)question the maximum score/weight is mentioned between brackets. The sum of these weights is 105 in total. However, only your best-answered questions, together giving a maximum score of 75, will be taken into account. **Hence, arbitrary sub-questions, together worth 30 points, should be skipped.** The final grade (scaled 1 through 10; $fail < 6.0 \leq pass$) will be determined for 75% by this written exam and for 25% by the three theory and three computer assignments (if these were handed in before their respective deadlines).

The grades will become available within 15 working days and will be announced by the FEE student administration office. Individual participants may inspect their results by making an appointment (preferably by E-mail) with the responsible professor.

Note that your handwriting should be clear, your notation consistent, and all your answers should be motivated. Below, we adopt the notation as used in Davidson & MacKinnon (2004), but we do not use bold-face for vectors and matrices, and use $'$ for transpose.

1. Consider the AR(1) model $y_t = \beta y_{t-1} + u_t$, where $|\beta| < 1$ and $u_t = \varepsilon_t + \alpha\varepsilon_{t-2}$ with $\varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2)$. Observations are available on $y_0, ..., y_n$, whereas y_0 is independent of $\varepsilon_1, ..., \varepsilon_n$.
   (a) {5} Show that $E(y_{t-1}u_t) = 0$ when $\alpha = 0$. What important consequence does this have for the OLS estimator of β?
   (b) {5} Clarify whether the OLS estimator of β is biased or unbiased when $\alpha = 0$.
   (c) {5} From now on we assume that α may be different from zero. What is then the most important property of the OLS estimator of β? Why?
   (d) {5} Would it be a good idea to use both y_{t-2} and y_{t-3} as instruments and estimate β by applying (generalized) IV to the AR(1) model? Why (not)?

2. A bank is interested in understanding the behavior of its clients regarding investing in risky assets. Data are collected and a model is formulated of the following form

$$y_i^* = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + u_i,$$

where x_{i2} denotes the age of client i and x_{i3} and x_{i4} her/his income and other savings respectively. The amount (s)he invested in risky assets is given by the limited dependent variable y_i :

$$\begin{cases} y_i = y_i^* & \text{if } y_i^* > 0 \\ y_i = 0 & \text{otherwise.} \end{cases}$$

It is assumed that $u_i \sim \text{NID}(0, \sigma^2)$ and is independent of all explanatory variables. Whether or not a client is investing in risky assets can be indicated by the discrete variable d_i :

$$\begin{cases} d_i = 1 & \text{if } y_i^* > 0 \\ d_i = 0 & \text{otherwise.} \end{cases}$$

(a) {5} Derive the probability that $d_i = 1$ as a function of the explanatory variables and the parameters.

(b) {5} Write down the log likelihood function for the binary (probit) model.

(c) {10} Derive the first-order conditions for a maximum of the log likelihood.

(d) {5} Now write down the log likelihood function for the censored (Tobit) model.

(e) {5} Indicate how you would test the hypothesis that age does not affect the amount of investment in risky assets in the Tobit model by a likelihood ratio test. What are the hypotheses? When would you (not) reject?

(f) {5} This hypothesis could also be tested in the probit model. Which of the two testing procedures would you prefer, and why?

3. Consider the system of two linear equations

$$y_i = X_i\beta_i + u_i, \qquad i = 1,2$$

where y_i, X_i and u_i are $n \times 1$ vectors and β_i is scalar. The full system can be written as

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

We assume that

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim \left(0, \begin{bmatrix} \sigma_{11}I_n & \sigma_{12}I_n \\ \sigma_{21}I_n & \sigma_{22}I_n \end{bmatrix}\right).$$

(a) {10+10} Prove for two particular situations that applying GLS to the full system is equivalent to applying OLS to the two separate equations.

(b) {10} Describe step by step and explicitly how you would implement FGLS estimation when the σ_{ij} parameters are unknown for $i, j = 1, 2$.

4. For finding the ML (maximum likelihood) estimator $\hat{\theta}_{ML}$ of a $k \times 1$ parameter vector θ we need to maximize the scalar log likelihood function $\ell(\theta)$ which depends on the observed data and the unknown parameters θ. Newton's method makes use of the $k \times 1$ gradient vector $g(\theta) \equiv \frac{\partial \ell(\theta)}{\delta \theta}$, the $k \times k$ Hessian matrix $H(\theta) \equiv \frac{\partial g(\theta)}{\delta \theta'}$ and the second order Taylor approximation

$$\ell(\theta) \approx \ell(\theta_{(j)}) + g(\theta_{(j)})'(\theta - \theta_{(j)}) + \frac{1}{2}(\theta - \theta_{(j)})'H(\theta_{(j)})(\theta - \theta_{(j)}),$$

where $\theta_{(j)}$ denotes the value of the vector of estimates at step j of the Newton algorithm.

(a) {10} Differentiate the 2nd order Taylor approximation formula with respect to θ, evaluate the result in $\hat{\theta}_{ML}$, make use of the value of $g(\hat{\theta}_{ML})$, and then show that this yields the fundamental equation for Newton's method

$$\theta_{(j+1)} = \theta_{(j)} - [H(\theta_{(j)})]^{-1} g(\theta_{(j)}).$$

(b) {10} Show that $\ell(\theta_{(j+1)}) > \ell(\theta_{(j)})$, i.e. the iteration step yields an increased log likelihood, when $H(\theta_{(j)})$ is negative definite as long as $g(\theta_{(j)}) \neq 0$. Hint: When A is negative definite then A^{-1} is negative definite too.

Success!