Exam Advanced Econometrics 2 (FEE-UvA: 2004/2005)
resit of 6 July, 2005; 9.30-11.30 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 3 weeks and will be announced by the FEE student administration office. Individual participants may inspect their results by making an appointment (preferably by email) with the responsible professor.

Note that your handwriting should be clear, your notation consistent, and all your answers should always 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 specific nonlinear autoregressive model

$$y_t = x_t(\beta) + u_t = \beta_1 \exp(\beta_2 y_{t-1}) + u_t, \qquad t = 1, ..., n$$

where the u_t are NID$(0, \sigma^2)$ and thus have density $(2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2}u_t^2/\sigma^2)$. Let the $k \times 1$ vector θ contain the $k = 3$ parameters β_1, β_2 and σ^2. We consider a few specific and some more general results.

(a) Argue why the density of the n-element sample $y = (y_1, ..., y_n)'$ is given here by

$$f(y \mid y_0, \theta) = \prod\nolimits_{t=1}^{n} f(y_t \mid y_{t-1}, \theta).$$

(b) Derive $f(y_t \mid y_{t-1}, \theta)$ and the contributions $\ell_t(y_t, y_{t-1}, \theta)$ to the loglikelihood.

(c) Derive for $i = 1, ..., 3$ the elements $G_{ti}(y_t, y_{t-1}, \theta) = \partial \ell_t(y_t, y_{t-1}, \theta)/\partial\theta_i$ of the $n \times k$ matrix $G(y, y_0, \theta)$.

(d) How are the columns of the matrix $G(y, y_0, \theta)$ related to the elements of the score vector $g(y, y_0, \theta)$?

(e) Indicate why and when $E_\theta[G_{ti}(y_t, y_{t-1}, \theta) \mid y_{t-1}] = 0$, and show (step by step) that this implies $E_\theta[g(y, y_0, \theta)] = 0$.

(f) Show that, for $t \neq s$, $E_\theta[G_{ti}(y_t, y_{t-1}, \theta)G_{si}(y_s, y_{s-1}, \theta)] = 0$.

(g) Let I_t be the $k \times k$ matrix with typical element $E_\theta[G_{ti}(y_t, y_{t-1}, \theta)G_{tj}(y_t, y_{t-1}, \theta)]$ for $i, j = 1, ..., k$. Show that

$$\sum\nolimits_{t=1}^{n} I_t = E_\theta[g(y, y_0, \theta)g(y, y_0, \theta)'].$$

(h) {10} Indicate which of the various results obtained above and what further theorems play a role in obtaining a fundamental result on the distribution of $g(y, y_0, \theta_0)$, where θ_0 is the true value of θ.

2. We consider model selection of nested and nonnested models.

(a) What is the principle difference between nested and nonnested hypotheses testing?

(b) Why is the maximized loglikelihood not a very appropriate criterion for the selection of a series of nested models? What type of correction is called for?

(c) Let two nonnested models H_1 and H_2 be characterized by the loglikelihood functions $\ell^{(j)}(\theta^{(j)}) = \sum_{t=1}^{n} \ell_t^{(j)}(\theta^{(j)})$ for $j = 1, 2$. Suppose that the true DGP μ belongs to H_1 and not to H_2. Show by using Jensen's inequality, which implies for a regular random variable v the inequality $E_\mu(\log v) < \log E_\mu(v)$, that under μ

$$\operatorname*{plim}_{n\to\infty} \frac{1}{n}\left[\ell^{(1)}(\hat{\theta}^{(1)}) - \ell^{(2)}(\hat{\theta}^{(2)})\right] > 0,$$

where $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$ are the MLEs of the two models. Hint: Using $\ell_t^{(j)}(\theta^{(j)}) = \log f_t^{(j)}(\theta^{(j)})$, show that for any $\theta^{(2)}$ we have $E_\mu \log[f_t^{(2)}(\theta^{(2)})/f_t^{(1)}(\theta_0^{(1)})] < \log E_\mu[f_t^{(2)}(\theta^{(2)})/f_t^{(1)}(\theta_0^{(1)})]$, where $\theta_0^{(1)}$ denotes the true value under μ. Evaluate the latter expression and use the implications for $\theta^{(2)} = \hat{\theta}^{(2)}$.

(d) What do the results under (b) and (c) imply for the general model selection criterion: choose the model with the largest maximized loglikelihood?

3. Consider the nonlinear regression model $y_t = x_t(\beta) + u_t$ with $u_t = \sigma v_t$, $v_t \sim \text{NID}(0, 1)$ and $x_t(\beta)$ nonrandom. So, v_t has PDF $\phi(v_t) = (2\pi)^{-1/2}\exp(-\frac{1}{2}v_t^2)$ and CDF $\Phi(v_t)$ for $t = 1, ..., n$.

(a) Express the density of y_t in the parameters β and σ, the data y_t and regression function $x_t(\beta)$ upon using $\phi(\cdot)$.

(b) Let y_t be observed for $y_t < c$ only. Express $\Pr(y_t < c)$ in $\Phi(\cdot)$.

(c) Let $\tilde{y}_t = y_t$ if $y_t < c$ and consider this truncated sample. Derive an expression for the PDF of $\tilde{y}_t$ in terms of β, σ, $\tilde{y}_t$, $x_t(\beta)$ and using the functions $\Phi(\cdot)$ and $\phi(\cdot)$.

(d) Obtain the loglikelihood of the truncated sample, and show that for $c \to +\infty$ it tends to the loglikelihood of the untruncated sample.

Success!