Exam Advanced Econometrics 1 (FEB-UvA: 20??/20??)
October ??, 20??; 9.00-11.00 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.

The grades will become available within three weeks and will be announced by the FEB 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 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. Let $\hat{\theta}$ be an estimator (obtained from a sample of size n) of a $p \times 1$ vector of parameters θ with true value θ_0. Let the expected squared error of $\hat{\theta}$ exist and be defined as $ESE(\hat{\theta}) \equiv E[(\hat{\theta}-\theta_0)(\hat{\theta}-\theta_0)']$, whereas the bias of $\hat{\theta}$ can be indicated by $B(\hat{\theta}) \equiv E(\hat{\theta}-\theta_0)$.
   (a) {5} Derive $ESE(\hat{\theta}) = Var(\hat{\theta}) + B(\hat{\theta})B(\hat{\theta})'$.
   (b) {10} Let $\text{plim}\, n^{1/2}(\hat{\theta}-\theta_0) = v$, with $v \sim N(0, V^{\infty}(\hat{\theta}))$ and $V^{\infty}(\hat{\theta})$ finite and nonsingular. Derive what this (in regular cases) does imply for $\text{plim}\, nESE(\hat{\theta})$.
   (c) {10} How would you construct a confidence set for θ_1, where $\theta = (\theta_1' \; \theta_2')'$ with θ_1 a $p_1 \times 1$ vector $(0 < p_1 < p)$. Introduce explicitly any further statistics, quantities and assumptions that are involved, and formulate formally the major property of the resulting confidence set.

2. Time-series have been observed on variables y_t and x_t, $t = 1, ..., n$. Consider the following regression models

$$y = \beta_1 \iota + \beta_2 x + u, \qquad (1)$$

$$Dy = \beta_2 Dx + Du, \qquad (2)$$

where ι denotes an $n \times 1$ vector of ones and the $(n-1) \times n$ matrix D is defined as:

$$D = \begin{pmatrix} -1 & 1 & 0 & \cdots & \cdot & \cdot & 0 \\ 0 & -1 & 1 & & & & \cdot \\ \vdots & & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & & \vdots \\ \cdot & & & & -1 & 1 & 0 \\ 0 & \cdot & \cdot & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

In other words, D transforms into first differences. For example, the $n-1$ elements of Dy are $y_t - y_{t-1}$, $t = 2, ..., n$.

(a) {10} Derive algebraically (hence, do not simply refer to the Frisch-Waugh-Lovell theorem) that the OLS estimator of β_2 in (1) is given by $\hat{\beta}_2 = x'Ay/x'Ax$, with $A = I - \frac{1}{n}\iota\iota'$.

(b) {5} As usual for full column rank matrices X, we denote by $P_X \equiv X(X'X)^{-1}X'$ the matrix that projects on the space spanned by the columns of X. Show (without applying direct matrix multiplication) that $I - P_\iota = P_{D'}$. Hence, demonstrate for $A = I - P_\iota$ that it projects on the space spanned by the columns of D'.

(c) {5} Assume from now on that $E(u) = 0$ and $E(uu') = \sigma_0^2 I$. Derive the variance matrix of the transformed error term Du. Does Du involve heteroskedasticity and/or autocorrelation, and if so: of what form/order?

(d) {5} Demonstrate now that the Generalized Least Squares (GLS) estimator of β_2 in model (2) is equal to the OLS estimator of β_2 in (1).

(e) {10} Suppose $x_t = y_{t-1}$ (hence, y_0 has been observed too). Is the OLS estimator of model (2) consistent now? Why (not)? What further assumption(s) did you make?

(f) {5} Is when $x_t = y_{t-1}$ the GLS estimator of model (2) consistent? Why (not)?

3. Consider the following regression models

$$y_t = \delta z_{t1} + (1-\delta) z_{t2} - \frac{\rho}{2}\delta(1-\delta)\left(z_{t1} - z_{t2}\right)^2 + u_t, \qquad \text{(NL)}$$

and

$$\begin{aligned} y_t &= Z_t^* \pi + u_t \\ &= \pi_1 z_{t1} + \pi_2 z_{t2} + \pi_3 \left(z_{t1} - z_{t2}\right)^2 + u_t, \end{aligned} \qquad \text{(L)}$$

where we assume that $Z_t = (z_{t1}\ z_{t2})$ contains exogenous regressors and $u_t \sim IID(0, \sigma^2)$.

(a) {5} Write down the relationships between the parameters δ and ρ from model (NL) and π_1, π_2 and π_3 from (L). Which restriction(s) have been imposed?

(b) {5} The asymptotic distribution of the OLS estimator $\hat{\pi}$ of the parameter π of model (L) with true value $\pi_0 = (\pi_{10}, \pi_{20}, \pi_{30})'$ is given by $n^{1/2}(\hat{\pi} - \pi_0) \overset{a}{\sim} N(0, \sigma_0^2 S_{Z^{*\prime}Z^*}^{-1})$. Give consistent estimators of δ and ρ based on the OLS estimator $\hat{\pi}$. Are these estimators unique? Why (not)?

(c) {10} Derive the asymptotic distribution of these particular estimators $(\hat{\delta}, \hat{\rho})'$ in terms of (characteristics of) $\hat{\pi}$.

(d) {10} Consider the Gauss-Newton estimation procedure to obtain NLS estimates of δ and ρ. Suppose we start with the initial estimates $\beta_{(0)}$ obtained in (b). Indicate the essential details of the next iteration step, i.e. how is $\beta_{(1)}$ obtained?

(e) {5} What is special about $\beta_{(1)}$ for this particular model: Does it obey the restriction(s) obtained in (a)? Is it asymptotically efficient? Why (not)?

(f) {5} How can you obtain an exact test for $H_0 : \rho = 0$? Do you need extra assumptions? Which, and why?

Success!