Example Exam B Advanced Econometrics 2 (FEB-UvA)
00 January, 200?; 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.
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 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 (and occasionally for derivative too).

1. A researcher runs by NLS the regressions (1) $y_t = x_t(\beta) + u_t$ and (2) $y_t/w_t = x_t(\beta^*)/w_t + u_t^*$, where w_t is an observed strictly positive scalar variable and $E(u_t \mid x_t(\beta), w_t) = 0$, $\forall \beta$ and $t = 1, ..., n$. For $\varepsilon_t \sim \text{NID}(0,1)$ the PDF of ε_t is $f_{\varepsilon_t}(\varepsilon_t) = (2\pi)^{-1/2} \exp(-\varepsilon_t^2/2)$.

   (a) {5} Are the models (1) and (2) nested or nonnested? Why?

   (b) {5} Show that, assuming $u_t \sim \text{NID}(0, \sigma^2)$, the ML estimator of β is equivalent to the NLS estimator.

   (c) {10} Assuming $u_t \sim \text{NID}(0, \sigma^2)$ and $u_t^* \sim \text{NID}(0, \sigma^{*2})$ respectively, how would you compare the fit of these two models? Indicate what would be problematic when simply comparing the sum of squared residuals of the two NLS regressions.

2. We consider a model with sample selectivity, where $y_t^\circ = X_t\beta + u_t$ and $z_t^\circ = W_t\gamma + v_t$ with

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} \sim \text{NID}\left(0, \begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix}\right),$$

   for $t = 1, ..., n$. Both sets of regressor variables X_t (which is $1 \times k$) and W_t (which is $1 \times l$) are predetermined and have been observed. With respect to the two dependent variables one in fact observes z_t, where $z_t = 1$ if $z_t^\circ > 0$ and $z_t = 0$ otherwise, and y_t° is observed only for those t for which $z_t = 1$. Hence, we have $y_t = y_t^\circ$ if $z_t^\circ > 0$ and otherwise y_t is not available. This implies that the log likelihood can be written as

$$\ell(\beta, \gamma, \rho, \sigma) = \sum_{t=1}^{n} \{I(z_t = 0) \log [\Pr(z_t = 0)] + I(z_t = 1) \log [\Pr(z_t = 1 \mid y_t^\circ) f(y_t^\circ)]\},$$

   where $I(\cdot)$ is the indicator function and $f(y_t^\circ)$ the density of y_t°.

   (a) {5} Express $\log[\Pr(z_t = 0)]$ in Φ, the CDF of the standard normal distribution, and $W_t\gamma$.

   (b) {10} Show that v_t can be decomposed in $v_t = \phi u_t + \varepsilon_t$ such that u_t and ε_t are independent. Express ϕ in the parameters ρ and σ and derive the distribution of ε_t.

(c) {5} Next show that the above decomposition can be used to express z_t° in terms of the parameters $(\beta, \gamma, \rho, \sigma)$, the regressors X_t and W_t, the variable y_t° and ε_t.

(d) {5} Use the above to express $\Pr(z_t = 1 \mid y_t^\circ)$ in Φ and in unknown parameters and the observed variables W_t, X_t and y_t only.

(e) {10} Now collect the above results such that $\ell(\beta, \gamma, \rho, \sigma)$ can be evaluated for given argument $(\beta, \gamma, \rho, \sigma)$, and from this expression explain what will be the consequences when maximizing this function if the true value of $\rho = 0$ and this value is, or is not, imposed.

3. Consider the linear regression model $y = X\beta + u$, where there are n observations and k regressors whereas $E(u) = 0$ and $\text{Var}(u) = \Omega$. We have a $n \times l$ matrix W of instrumental variables with $n > l \geq k$, and with t-th row W_t such that $E(u_t \mid W_t) = 0$, $E(u_t u_s \mid W_t, W_s) = (\Omega_0)_{t,s}$, $(t, s = 1, ..., n)$ and $\text{plim}\, n^{-1/2} W'u \sim \text{N}(0, \text{plim}\, \frac{1}{n} W'\Omega_0 W)$, where Ω_0 is the true value of Ω. The true value of β is β_0.

   (a) {10} Consider the criterion function $(y - X\beta)'WW'(y - X\beta)$. Derive the expression for the coefficient estimator that minimizes this criterion function and show that this estimator (call it $\hat{\beta}_I$) is consistent (mention any further reasonable assumptions that you need to make).

   (b) {5} Derive the asymptotic distribution of estimator $\hat{\beta}_I$.

   (c) {5} Let the matrix $W^* = WA$, with A a nonsingular $l \times l$ transformation matrix, be such that W^* contains valid instruments, i.e. $\text{plim}\, n^{-1/2} W^{*\prime} u \sim \text{N}(0, \text{plim}\, \frac{1}{n} W^{*\prime} \Omega_0 W^*)$ with $\text{plim}\, \frac{1}{n} W^{*\prime} \Omega_0 W^*$ finite and nonsingular. Consider the estimator $\hat{\beta}_A$ obtained by minimizing $(y - X\beta)'WAA'W'(y - X\beta)$. Derive its asymptotic distribution, stressing its dependence on A.

   (d) {10} Indicate why it is plausible that the asymptotic variance of $\hat{\beta}_A$ is minimized by choosing A such that $\text{plim}\, AA' = (\text{plim}\, \frac{1}{n} W'\Omega_0 W)^{-1}$ and characterize this efficient estimator.

4. Consider the three equations simultaneous model

$$
\begin{aligned}
y_{t1} &= \theta_1 + \theta_2 w_{t2} + \theta_3 w_{t3} + \theta_4 y_{t2} + u_{t1} \\
y_{t2} &= \theta_5 + \theta_6 w_{t2} + \theta_7 w_{t4} + \theta_8 y_{t1} + u_{t2} \\
y_{t3} &= \theta_9 + \theta_{10} w_{t5} + \theta_{11} y_{t1} + \theta_{12} y_{t2} + u_{t3}
\end{aligned}
$$

   where y_{t1}, y_{t2} and y_{t3} are endogenous variables whereas the intercept and $w_{t2}, ..., w_{t5}$ are exogenous explanatory variables. All variables have been observed for $t = 1, ..., n$. About the disturbances we make the standard assumptions. All the coefficients $\theta_1, ..., \theta_{12}$ have unknown real values.

   (a) {5} This model can also be expressed as $Y\Gamma = WB + U$. Indicate clearly how all the variables and coefficients of this particular model are accommodated in this notation.

   (b) {5} Discuss the identification of the third equation of this structural model according to both the order and the rank condition.

   (c) {10} What are the consequences for the identification of the third equation if we did not know that w_{t2} does not occur (has zero coefficient) in the third equation? What particular true parameter values would be problematic then?

Success!