First Midterm Exam, Econ. 240B

Department of Economics
U.C. Berkeley
March 12, 2007

Instructions: You have 110 minutes to complete this exam. This is a 30 point exam; all subsections of all questions have equal weight (5 points each). This is a closed book exam, but one sheet of notes is permitted. All needed statistical tables are appended. Please make your answers elegant – that is, clear, concise, and correct.

1. **True/False/Explain** (15 points): For **three** of the following four statements below, determine whether it is correct, and, if correct, explain why. If not, state precisely why it is incorrect and give a modification which is correct. **Answer only three questions;** if you answer more, only the **first three** answers will count in your score.

A. Suppose that an IV regression of y_i on a scalar endogenous regressor x_{i1} and a vector x_{i2} of exogenous regressors, using an instrument vector z_i that includes the x_{i2} components, yields a coefficient on x_{i1} of 2.2. If, instead, x_{i1} is taken to be the dependent variable, and an IV fit of x_{i1} on y_i and x_{i2} is calculated using the same instruments z_i, then the IV estimate of the coefficient on y_i will be positive.

B. In the linear model with a lagged dependent variable, $y_t = x_t'\beta + \gamma y_{t-1} + \varepsilon_t$, suppose the error terms are $MA(1)$, i.e., $\varepsilon_t = u_t + \theta u_{t-1}$, where u_t is an i.i.d. sequence with zero mean, variance σ^2, and is independent of x_s for all t and s. For this model, the classical LS estimator will be inconsistent for β and γ when $|\gamma| < 1$, but an IV estimator using x_t and y_{t-2} as instrumental variables will consistently estimate these parameters.

C. For a balanced panel data regression model with individual fixed effects, $y_{it} = x_{it}'\beta + \alpha_i + \varepsilon_{it}$ – where the α_i are are not assumed to be uncorrelated with x_{it}, but the error terms ε_{it} are i.i.d. and independent of α_i and x_{it}, with $E(\varepsilon_{it}) = 0$ and $V(\varepsilon_{it}) = \sigma^2$ – suppose that only the number of time periods T tends to infinity, while the number of individuals N stays fixed. Then the "fixed effect" estimator for β will be consistent as $T \to \infty$ provided the regressors and individual indicator variables are not asymptotically multicollinear. Furthermore, if $\hat{\sigma}^2 = (NT)^{-1} \sum_i \sum_t (y_{it} - \hat{\alpha}_i - x_{it}'\hat{\beta}_{LS})^2$ is the (biased) LS estimator of σ^2, then the usual LS formulae for the standard errors of $\hat{\beta}_{LS}$ (replacing the unknown σ^2 by $\hat{\sigma}^2$) will be asymptotically valid.

D. By the so-called "Delta Method", if $\hat{\theta}$ is root-n consistent and asymptotically normal for a vector parameter θ_0, then the difference between the squared length of $\hat{\theta}$ and the squared length of θ_0, when multiplied by the square root of the sample size, will generally have a limiting normal distribution.

2. (5 points) Suppose a dependent variable y_i and two (scalar) regressors x_i and z_i satisfy a random coefficients model

$$y_i = \alpha_i + \beta_i x_i + \gamma_i z_i, \qquad i = 1, ..., N,$$

where the coefficients $(\alpha_i, \beta_i, \gamma_i)$ are assumed to be i.i.d. and independent of x_i and z_i. In this framework, under the null hypothesis $H_0 : Var(\beta_i) = 0 = Var(\gamma_i)$, the mean values $\beta = E(\beta_i)$ and $\gamma = E(\gamma_i)$ can be estimated by a least-squares regression of y_i; in turn, this null hypothesis can be tested using the R^2 from a least-squares regression of the squared LS residuals $\hat{e}_i^2 = (y_i - \hat{\alpha} - \hat{\beta} x_i - \hat{\gamma} z_i)^2$ on functions of the regressors.

Given a sample of size $N = 500$, derive the algebraic form of all of the regressors in this "squared residual regression", and give a numerical value for the critical value C for an (asymptotic) 5% test of homoskedasticity using the second-stage R^2. i.e., the value for which H_0 will be rejected if $R^2 > C$ with asymptotic size 5%.

3. (5 points) A feasible GLS fit of the generalized regression model with $K = 3$ regressors yields the estimates $\hat{\beta} = (2, -2, -1)$. where the GLS covariance matrix $V = \sigma^2 [X'\Omega^{-1}X]^{-1}$ is estimated as

$$\hat{V} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

using consistent estimators of σ^2 and Ω. The sample size $N = 403$ is large enough so that it is reasonable to assume a normal approximation holds for the GLS estimator.

Use these results to test the null hypothesis $H_0 : \beta_1^2 + \beta_2^2 + \beta_3^2 = 1$ at an asymptotic 5% level.

4. (5 points) If y_t is an $MA(1)$ process with zero mean, i.e., if

$$y_t = \varepsilon_t + \theta \varepsilon_{t-1}, \qquad \varepsilon_t \sim WN(\sigma^2),$$

and if $\gamma(s) = Cov(y_t, y_{t-s})$ is the autocovariance function and $\rho(s) = \gamma(s)/\gamma(0)$ is the autocorrelation function of $\{y_t\}$, show that

$$-1 < c^L \leq \rho(1) \leq c^U < 1,$$

i.e., the first autocorrelation is strictly bounded away from -1 and 1, by calculating the maximum and minimum values c^U and c^L of $\rho(1)$ over all possible θ.