

First Midterm Exam, Econ. 240B
Department of Economics
U.C. Berkeley
March 15, 2004

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. If the Generalized Regression model holds – that is, $E[y|X] = X\beta$, $V[y|X] = \sigma^2\Omega$, and X full rank with probability one – then the covariance matrix between Aitken’s Generalized LS estimator $\hat{\beta}_{GLS}$ (with known Ω matrix) and the classical LS estimator $\hat{\beta}_{LS}$ is equal to the variance matrix of the LS estimator.

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 have first-order serial correlation, i.e., $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$, 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 , but Aitken’s GLS estimator (for a known Ω matrix) will consistently estimate these parameters.

C. For a balanced panel data regression model with random individual effects, $y_{it} = x_{it}'\beta + \alpha_i + \varepsilon_{it}$ (where the α_i are independent of ε_{it} and x_{it} , and all error terms have mean zero, constant variance, and are serially independent across i and t), 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 $T \rightarrow \infty$, but the “random effects” GLS estimator is infeasible, since the joint covariance matrix of the error terms is not consistently estimable.

D. In a linear model with an intercept and two nonrandom, nonconstant regressors, and with sample size $N = 200$, it is suspected that a “random coefficients” model applies, i.e., that the intercept term and two slope coefficients are jointly random across individuals, independent of the regressors. If the squared values of the LS residuals from this model are themselves fit to a quadratic function of the regressors, and if the R^2 from this second-step regression equals 0.06, the null hypothesis of no heteroskedasticity should be rejected at an approximate 5% level.

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

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

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 $H_0 : \theta = 1$ against a two-sided alternative at an asymptotic 5% level, where

$$\theta \equiv g(\beta) = \|\beta\| \equiv (\beta_1^2 + \beta_2^2 + \beta_3^2)^{1/2}.$$

3. (10 points) Consider the estimation of two scalar coefficients, β_1 and β_2 , in the linear equation

$$y = x_1\beta_1 + x_2\beta_2 + \varepsilon,$$

where y , x_1 , and x_2 are observable N -dimensional random vectors. In addition, two N -dimensional vectors of instrumental variables, z_1 and z_2 , are available. In a sample of size $N = 227$, the following matrix of cross-products of the variables is observed:

$$\begin{bmatrix} y'y & y'x_1 & y'x_2 & y'z_1 & y'z_2 \\ x_1'y & x_1'x_1 & x_1'x_2 & x_1'z_1 & x_1'z_2 \\ x_2'y & x_2'x_1 & x_2'x_2 & x_2'z_1 & x_2'z_2 \\ z_1'y & z_1'x_1 & z_1'x_2 & z_1'z_1 & z_1'z_2 \\ z_2'y & z_2'x_1 & z_2'x_2 & z_2'z_1 & z_2'z_2 \end{bmatrix} = \begin{bmatrix} 22 & -11 & 10 & 8 & 8 \\ -11 & 21 & 10 & -8 & -8 \\ 10 & 10 & 20 & -2 & 0 \\ 8 & -8 & -2 & 6 & 4 \\ 8 & -8 & 0 & 4 & 6 \end{bmatrix}$$

A. For these data, calculate the classical LS estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ of the unknown regression coefficients, and compute the instrumental variables estimators $\tilde{\beta}_1$ and $\tilde{\beta}_2$ using z_1 and z_2 as instruments for x_1 and x_2 .

B. Suppose the error terms ε are independent of z_1 and z_2 , so that $V[\varepsilon|z_1, z_2] = \sigma^2 I$, i.e., ε has a scalar covariance matrix. If you had to conduct a test of $H_0 : \beta_2 = 1$ versus $H_0 : \beta_2 \neq 1$ at an asymptotic 5% level using the IV estimator, and were given a consistent estimator $\tilde{\sigma}^2$ of the unknown variance parameter σ^2 , how small would $\tilde{\sigma}^2$ have to be to reject H_0 ? That is, find the largest value of $\tilde{\sigma}^2$ for which you could (barely) reject the null hypothesis.