BOSTON COLLEGE
Department of Economics

EC 771B
Econometrics
Spring 2000

MIDTERM EXAMINATION
11 April 2000

Answer all questions. Total of 139 points. Partial credit given for partial answers.

1. (25 pts) Briefly define each of the following terms in one or two sentences.

a. Two-way ANOVA:*refers to the regression of* y *on two sets of dummy variables, or qualitative factors in categorical form*

b. Condition number of $(X'X)$: *not a problem*

c. Davidson-MacKinnon "J" test: *this is not Hansen's J test in GMM, but the test for non-nested models*

d. Asymmetry of specification error: *refers to the severe consequences of excluding relevant variables vs. milder consequences of including irrelevant variables in a regression*

e. Box-Pierce "Q" test: *not a problem*

2. (30 pts) Suppose that we have cross-sectional data on workers' earnings and their age, and want to estimate an earnings-age profile in which

$$E_i = \alpha + \beta A_i + \epsilon_i$$

but we believe that the profile may have a different slope for the age groups $A \leq 18$, $19 \leq A \leq 61$, $A \geq 62$.

a. Write down the single equation that could be used to estimate this set of earnings profiles. Define any variables that you need.

Define $D_{19} = 1$ *for* $A \geq 19$ *and* $D_{62} = 1$ *for* $A \geq 62$ *(sic). Then the single regression is*

$$E_i = \beta_1 + \beta_2 D_{19} + \beta_3 D_{62} + \beta_4 A_i + \beta_5 A_i D_{19} + \beta_6 A_i D_{62} + \epsilon_i$$

with i subscripts on the dummies suppressed. Note that regression on the dummies alone does not estimate $\partial E/\partial A$, but merely computes conditional means of E by age group.

b. The results of estimating this equation could be criticized as unusable in forecasting earnings due to discontinuities in the function. Show how the equation may be reestimated, allowing for different slopes for each age group, but yielding a continuous (not necessarily differentiable) function. How many distinct parameters are now being estimated?

To impose continuity at the two "knot points" and create a linear spline, two conditions are required:

$$\beta_1 + \beta_4 A_i\rfloor_{19} = (\beta_1 + \beta_2) + (\beta_4 + \beta_5)\, A_i\rfloor_{19}$$

$$(\beta_1 + \beta_2) + (\beta_4 + \beta_5)\, A_i\rfloor_{62} = (\beta_1 + \beta_2 + \beta_3) + (\beta_4 + \beta_5 + \beta_6)\, A_i\rfloor_{62}$$

which yield the linear restrictions $0 = \beta_2 + \beta_5 A_i\rfloor_{19}$ and $0 = \beta_3 + \beta_6 A_i\rfloor_{62}$. These two linear restrictions on the six parameters imply that four parameters are now being estimated in the continuous earnings profile.

c. Discuss how the original model would be modified, and how you would constrain it, to produce an earnings profile that was both continuous and (once) differentiable, taking account of the three age groups.

The original model should be modified with an A_i^2 term, with interacted dummies being created for the three age groups for both A and A^2. The unconstrained model would have nine parameters (three intercepts, three age terms, three age^2 terms). Linear restrictions would then be specified on those parameters to impose continuity (one restriction for each knot point) and first differentiability (one restriction for each knot point). Thus four constraints would be imposed on these nine parameters to yield an earnings-age profile as a quadratic spline, continuous and once differentiable.

3. (24 pts) Consider the simple regression $y_i = \beta x_i + \epsilon_i$, where y and x are scalars.

a. Prove that the estimator $\tilde{\beta} = \bar{y}/\bar{x}$ is an unbiased estimator of β.

$\tilde{\beta} = \bar{y}/\bar{x}$, so $\bar{y} = \beta\bar{x} + \bar{\epsilon}$ and $\bar{y}/\bar{x} = \beta + \bar{\epsilon}/\bar{x}$. Applying expectations, $E\left[\tilde{\beta}\right] = \beta$.

b. Derive the least squares estimator of β, $\hat{\beta}$.

min $e'e = (y - bx)'(y - bx) = y'y - 2b'x'y + b'x'xb$. FOC yields $\hat{b} = (x'x)^{-1}x'y$ with sampling variance (needed in part c) of $E\left((x'x)^{-1}x'y - \beta\right)^2 = \sigma_\epsilon^2(x'x)^{-1}$.

c, Demonstrate that the least squares estimator $\hat{\beta}$ has a smaller sampling variance than that of $\tilde{\beta}$.

$E(\bar{y}/\bar{x} - \beta)^2 = E\left(\frac{\beta\bar{x}+\bar{\epsilon}}{\bar{x}} - \beta\right)^2 = E\left(\frac{\bar{\epsilon}}{\bar{x}}\right)^2 = \frac{\sigma_\epsilon^2}{n\bar{x}}$.

Since $x'x = \sum x_i^2 = \sum(x_i - \bar{x})^2 + n\bar{x}^2$ the ratio of sampling variances is

$$\frac{Var\left(\tilde{\beta}\right)}{Var(\hat{b})} = \frac{\sigma_\epsilon^2/n\bar{x}^2}{\sigma_\epsilon^2/x'x} = \frac{\sum(x_i - \bar{x})^2 + n\bar{x}^2}{n\bar{x}^2} > 1.$$

4. (15 pts) In the classical linear regression model $y = X\beta + \epsilon$, what is the covariance matrix of the GLS estimator of β ($\hat{\beta}$) and the difference between it and the OLS estimator of β, $\hat{b}$, that is,

$$Cov\left[\hat{\beta}, \hat{\beta} - \hat{b}\right]$$

where $\hat{\beta} = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}y)$ and $\hat{b} = (X'X)^{-1}(X'y)$?

$\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\epsilon$, while $\hat{b} = (X'X)^{-1}X'\epsilon$.

Thus $\hat{\beta} - \hat{b} = \left[(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} - (X'X)^{-1}X'\right]\epsilon$ with $E\left[\hat{\beta} - \hat{b}\right] = 0$, since both are unbiased estimators. $Cov\left[\hat{\beta}, \hat{\beta} - \hat{b}\right] = E\left[\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \hat{b}\right)'\right] =$

.

$$E\left[\left(\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}\epsilon\right)\epsilon'\left[\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1} - (X'X)^{-1}X'\right]'\right] =$$

$$\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}\sigma^2\Omega\left[\Omega^{-1}X\left(X'\Omega^{-1}X\right)^{-1} - X(X'X)^{-1}\right] =$$

$$\sigma^2\left[\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}\Omega\Omega^{-1}X\left(X'\Omega^{-1}X\right)^{-1} - \left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}\Omega X(X'X)^{-1}\right] = 0.$$

5. (15 pts) Does first differencing of time series data reduce autocorrelation? Consider the model

$$y_t = \beta' x_t + \epsilon_t,$$

$$\epsilon_t = \rho \epsilon_{t-1} + v_t$$

Compare the autocorrelation of ϵ in the original model with that of u_t in the model

$$y_t - y_{t-1} = \beta' (x_t - x_{t-1}) + u_t,$$

$$u_t = \epsilon_t - \epsilon_{t-1}.$$

For the AR(1) model, the autocorrelation is ρ. *For the first-differenced process,* $Var(u_t) = 2Var(\epsilon) - 2Cov(\epsilon_t, \epsilon_{t-1}) = 2\sigma_\nu^2 \left[\frac{1}{1-\rho^2} - \frac{\rho}{1-\rho^2}\right] = 2\sigma_\nu^2 / (1+\rho)$.

So $Cov(u_t, u_{t-1}) = 2Cov(\epsilon_t, \epsilon_{t-1}) - Var(\epsilon_t) - Cov(\epsilon_t, \epsilon_{t-1}) = \frac{\sigma_\nu^2}{(1-\rho^2)} [2\rho - 1 - \rho^2] = \sigma_\nu^2 \frac{\rho - 1}{\rho + 1}$,

so the autocorrelation of the u *process is* $\frac{Cov(u_t, u_{t-1})}{Var(u_t)} = \frac{\rho-1}{2}$. *This will be less than* ρ *for* $|\rho| > \frac{1}{3}$, *so if the* ϵ *process has a greater degree of autocorrelation than that, the first differenced process will have a lower degree of autocorrelation.*

6. (15 pts) Consider a generalization of the random effects model for use with panel data, in which

$$y_{it} = \alpha + \beta' x_{it} + \epsilon_{it} + \mu_i + \lambda_t$$

where

$$E[\epsilon_{it}] = E[\mu_i] = E[\lambda_t] = 0,$$

$$E[\epsilon_{it}\mu_i] = E[\epsilon_{it}\lambda_s] = E[\lambda_t\mu_i] = 0, \forall i, j, t, s : (CORRECTED)$$

$$Var[\epsilon_{it}] = \sigma_\epsilon^2, Cov[\epsilon_{it}, \epsilon_{js}] = 0, \forall i, j, t, s$$

$$Var[\mu_i] = \sigma_\mu^2, Cov[\mu_i, \mu_j] = 0, \forall i, j$$

$$Var[\lambda_t] = \sigma_\lambda^2, Cov[\lambda_t, \lambda_s] = 0, \forall t, s$$

Write out the full covariance matrix for the composite error term for a data set with $N = 2$ and $T = 2$.

$$\begin{array}{cccc}
\sigma_\epsilon^2 + \sigma_\mu^2 + \sigma_\lambda^2 & \sigma_\mu^2 & \sigma_\lambda^2 & 0 \\
\sigma_\mu^2 & \sigma_\epsilon^2 + \sigma_\mu^2 + \sigma_\lambda^2 & 0 & \sigma_\lambda^2 \\
\sigma_\lambda^2 & 0 & \sigma_\epsilon^2 + \sigma_\mu^2 + \sigma_\lambda^2 & \sigma_\mu^2 \\
0 & \sigma_\lambda^2 & \sigma_\mu^2 & \sigma_\epsilon^2 + \sigma_\mu^2 + \sigma_\lambda^2
\end{array}$$

with t varying within i (that is, the first 2x2 block refers to the errors from unit 1 for time periods 1 and 2).

7. (15 pts) Consider the multivariate regression model

$$y_1 = \alpha_1 + \beta_1 x + \epsilon_1$$

$$y_2 = \alpha_2 + \epsilon_2$$

$$y_3 = \alpha_3 + \epsilon_3$$

Suppose that we know that $y_2 + y_3 \equiv 1.0$ at every observation in the data. Show that these three equations cannot be estimated as a system of seemingly unrelated regressions, and indicate why that is so.

OLS on these equations will yield $\alpha_1 = \bar{y}_1, \alpha_2 = \bar{y}_2, \alpha_3 = \bar{y}_3$, and $\hat{\beta}_1 = (x'x)^{-1} x'y_1$. The residual vectors are

$$e_{i1} = y_{i1} - \bar{y}_1 - \hat{\beta}_1 x_1$$

$$e_{i2} = y_{i12} - \bar{y}_2$$

$$e_{i3} = y_{i13} - \bar{y}_3$$

but if $y_2 + y_3 \equiv 1.0$ for each i, then $\bar{y}_2 + \bar{y}_3 = 1$ as well, and $e_{i2} + e_{i3} = y_{i12} + y_{i13} - (\bar{y}_2 + \bar{y}_3) = 0$ for every observation. The residual covariance matrix is computationally singular. SUR cannot be applied, since it involves inverting this matrix.