#37 P.S.

ECONOMETRICS
COMPREHENSIVE EXAM
SUMMER 2004

YOON/KUMBHAKAR

ANSWER **ANY FOUR** QUESTIONS

1. Consider the linear regression model $y = X_1\beta_1 + X_2\beta_2 + u$ (1)
where X_1 and X_2 are $n_1 \times k$ and X_2 is $n_2 \times k$ matrices.

(i) Show that the OLS estimator of β_2 from (1) is also the OLS estimator β_2 from regressing $\widetilde{X}_2$ on $\widetilde{Y}$ where

$$\widetilde{Y} = \left[I - X_1\left(X_1^{1} X_1\right)^{-1} X_1^{1}\right]y \text{ and } \widetilde{X}_2 = \left[I - X_1\left(X_1^{1} X_1\right)^{-1} X_1^{1}\right]X_2 \quad (2)$$

This result is knows as the Frisch-Waught-Lovell (FWL) Theorem.

(ii) Using the FWL Theorem, show that if $X_1 = l_n$ a vector of ones indicating the presence of the constant in the regression, and X_2 is a set of economic variables, then (a) $\hat{\beta}_{2,OLS}$ can be obtained by running $y_i - \bar{y}$ on the set of variables in X_2 expressed as deviations from their respective sample means. (b) The least squares estimate of the constant $\hat{\beta}_{1,OLS}$ can be retrieved as $\bar{y} - \bar{X}_2'\hat{\beta}_{2,OLS}$ where $\bar{X}_2' = l_n' X_2 / n$ is the vector of sample means of the independent variables in X_2.

(iii) Let $y = X\beta + D_i\gamma + u$ where y is $n \times 1$, X is $n \times k$ and D_i is a dummy variable that takes the value 1 for the i-th observation and 0 otherwise. Using the FWL Theorem, prove that the least squares estimates of β and γ from this regression are $\hat{\beta}_{OLS} = \left(X^{*'}X^{*}\right)^{-1} X^{*'}y^{*}$ and $\hat{\gamma}_{OLS} = y_i - x_i'\hat{\beta}_{OLS}$, where X^* denotes the X matrix without the i-th observation and y^* is the y vector without the i-th observation and (y_i, x_i') denotes the i-th observation on the dependent and independent variables. This means that $\hat{\gamma}_{OLS}$ is the forecasted OLS residual from the regression of y^* on X^* for the i-th observation which was essentially excluded from the regression by the inclusion of the dummy variable D_i.

2. Consider the following two sets of regressions:

$$y_1 = X_1\beta_1 + u_1 \text{ and } y_2 = X_2\beta_2 + u_2 \quad (1)$$

X_1 is $n_1 \times k$ and X_2 is $n_2 \times k$ with $n_1, n_2 > k$.

Write the equations in (1) as $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ (2)

and

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta_1 + \begin{bmatrix} 0 \\ X_2 \end{bmatrix} (\beta_2 - \beta_1) + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (3)$$

(a) Show that OLS on (2) yields OLS on each equation separately in (1). In other words, $\hat{\beta}_{1,OLS} = (X_1'X_1)^{-1} X_1'y_1$ and $\hat{\beta}_{2,OLS} = (X_2'X_2)^{-1} X_2'y_2$.

(b) Show that the residual sum of squares for equation (2) is given by RSS_1 + RSS_2, where RSS_i is the residual sum of squares from running y_i on X_i for i = 1,2.

(c) Show that the Chow F-statistic can be obtained from (3) by testing for the joint significance of $H_0 : \beta_2 - \beta_1 = 0$.

3. Consider a simple dynamic panel data model. Discuss the problems associated with estimating the model using OLS, Within and GLS estimators. Suggest some methods (discuss at least one in details) for obtaining consistent estimators.

4. Assume the standard linear model: $Y = X\beta + \epsilon$, where $V(\epsilon) = \sigma^2 I$; Y and ϵ are nx1; X is nxk; β is kx1. Denote the least squares (LS) estimator of β by b. Then, the LS residual vector, e, is defined by Y - Xb. (V stands for covariance matrix.)

1) Show that E(e) = 0, where 0 is an nx1 vector of zeroes.
2) Derive the variance of the LS residual vector e.
3) Consider a special case of k = 1 so that $X'X = \Sigma^n_{j=1} x_j^2$ which is 1x1. Assume that X is not an intercept variable and n is larger than 5. Find the covariance between e_1 and e_2.

5. Assume a heteroscedastic model: $Y = X\beta + \epsilon$, where $V(\epsilon) \neq \sigma^2 I$. Y and are nx1; X is nxk; ϵ is kx1.

1) Derive V(b) where b denotes the least squares estimator of β.
2) Write the estimator of V(b) suggested by H. White.

6. You want to estimate the effect of college education denoted by dummy variable C on earnings y where $y_i = \beta' x_i + \delta C_i + \varepsilon_i$. Assuming that C is correlated with ε, show how you would perform the estimation. Make assumptions if necessary.