Problem Set #2

Economics 240B
Spring 2006

Due February 27

PART I: "Theoretical" questions:

Turn in (correct) answers to the following exercises from Ruud's text:

Chapter 18: Exercises 18.1, 18.2, 18.5, 18.7.
Chapter 19: Exercises 19.3, 19.5.
Chapter 26: Exercise 26.7

PART II: "Empirical" questions:

A famous problem in economics (studied by Ernst Engel) is whether food expenditure is proportionally related to income. Defining

$$\begin{aligned} y_i &= \text{logarithm of food expenditure for individual } i\text{, and} \\ x_i &= \text{logarithm of income for individual } i, \end{aligned}$$

then null hypothesis of proportionality is equivalent to the null hypothesis that, if y_i and x_i are assumed to follow the standard (bivariate) linear regression model, then the true value of the regression coefficient for x_i equals one.

(a) Some of Engel's original data for English working-class families (circa 1900) are available in the text file ENGEL.TXT. Use the 235 observations on EXP (food expenditure) and INC (income) to test the null hypothesis of proportionality at the 5% level, assuming the error terms are i.i.d. and normally distributed.

(b) Carry out the same test as in part (a), but use the Eicker-White covariance matrix estimator instead of the usual normal-theory estimator. Do your conclusions change?

(c) Regress the squared values of the residuals against a constant term, x_i and x_i^2 and test the null hypothesis of homoskedastic error terms at an (approximate) significance level of 5%.

(d) Use the predicted values from the "squared residual regression" in part (b) above as weights in a weighted least squares regression of y_i on x_i , and perform the analogous test of proportionality, assuming that the implicit quadratic specification for heteroskedasticity in part (c) above is correct. Do the conclusions from part (a) above change?

(e) Now suppose the quadratic (in x_i) model for the conditional variance implicit in (c) is incorrect, so that the weighted least squares estimator of part (d) is not efficient. Carry out the test for proportionality using the WLS estimators of (d) and a covariance matrix estimator which is consistent when the model for the conditional variance is misspecified. As usual, compare your conclusions to the ones in part (a) above.

2. Suppose you had a two-equation linear model

$$y_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta}_j + \varepsilon_{ij}, \qquad i = 1, ..., N, \qquad j = 1, 2,$$

but you suspected the errors might be heteroskedastic for each equation, and the correlation in the errors across equations (but not across individuals) might vary across individuals, i.e.,

$$Cov(\varepsilon_{ij}, \varepsilon_{i'k}) = \begin{cases} 0 & \text{if } i \neq i'; \\ \sigma_{i,jk} & \text{if } i = i'. \end{cases}$$

(a) One special case of this setup would be a random coefficient model

$$y_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta}_{ij}, \qquad i = 1, ..., N, \qquad j = 1, 2,$$

with $\boldsymbol{\beta}_{ij}$ random across i and independent of the regressors $\mathbf{x}_{i1}$ and $\mathbf{x}_{i2}$,

$$\begin{aligned} E[\boldsymbol{\beta}_{ij}] &\equiv \boldsymbol{\beta}_j, \\ C[\boldsymbol{\beta}_{ij}, \boldsymbol{\beta}_{i'k}] &= \begin{cases} \mathbf{0} & \text{if } i \neq i'; \\ \boldsymbol{\Gamma}_{jk} & \text{if } i = i', \end{cases} \end{aligned}$$

so that the original model holds with

$$\varepsilon_{ij} \equiv \mathbf{x}'_{ij}(\boldsymbol{\beta}_{ij} - \boldsymbol{\beta}_j).$$

Describe how you could construct Feasible GLS estimators of the regression coefficients $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ for this model.

(b) If you were unwilling to impose the random coefficient structure on this two-equation system, but wanted to test an r-dimensional nonlinear hypothesis

$$H_0 : g(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \mathbf{0}$$

on the regression coefficients for both equations, describe how you could construct an asymptotically-valid test for this hypothesis using the classical LS estimator

$$\hat{\boldsymbol{\beta}}_{LS} \equiv \begin{pmatrix} \hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2 \end{pmatrix} = \begin{pmatrix} (\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}_1\mathbf{y}_1 \\ (\mathbf{X}'_2\mathbf{X}_2)^{-1}\mathbf{X}_2\mathbf{y}_2 \end{pmatrix}.$$

Give algebraic expressions for the components of your test statistic and precisely state the critical region for the test.