Econometrics 710
Midterm Exam, March 25, 2004

1. Take the homoskedastic model

$$\begin{aligned} y_i &= x_{1i}'\beta_1 + x_{2i}'\beta_2 + e_i \\ E\left(e_i \mid x_{1i}, x_{2i}\right) &= 0 \\ E\left(e_i^2 \mid x_{1i}, x_{2i}\right) &= \sigma^2 \\ E\left(x_{2i} \mid x_{1i}\right) &= \Gamma x_{1i} \\ \Gamma &\neq 0 \end{aligned}$$

Suppose the parameter β_1 is of interest, and suppose that it is estimated by a regression of y_i on x_{1i} only. We know that the exclusion of x_{2i} makes the estimator biased and inconsistent for β_1. It also changes the equation error. Our question is: what is the effect on the homoskedasticity property of the induced equation error? Does the exclusion of x_{2i} induce heteroskedasticity or not? Be specific.

2. The model is

$$\begin{aligned} y_i &= x_i'\beta + e_i \\ E\left(e_i \mid x_i\right) &= 0 \\ E\left(e_i^2 \mid x_i\right) &= \sigma_i^2 \\ \Omega &= diag(\sigma_1^2, ..., \sigma_n^2). \end{aligned}$$

The parameter β is estimated both by OLS $\hat{\beta} = \left(X'X\right)^{-1} X'Y$ and GLS $\tilde{\beta} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y$. Let $\hat{e}_i = y_i = x_i'\hat{\beta}$ and $\tilde{e}_i = y_i = x_i'\tilde{\beta}$ denote the residuals and $\hat{R}^2 = 1 - \hat{e}'\hat{e}/(y^{*\prime}y^*)$ and $\tilde{R}^2 = 1 - \tilde{e}'\tilde{e}/(y^{*\prime}y^*)$ where $y = y - \overline{y}$ denote the equation R^2. If the error e_i is truly heteroskedastic will $\hat{R}^2$ or $\tilde{R}^2$ be smaller?

3. Take the model

$$\begin{aligned} y_i &= x_{1i}'\beta_1 + x_{2i}'\beta_2 + e_i \\ E\left(x_i e_i\right) &= 0 \end{aligned}$$

where β_1 and β_2 are each $k \times 1$. How would you test the joint hypothesis that the ratio of each element of β_1 and β_2 is one? That is, if $k = 1$, $H_0 : \beta_1/\beta_2 = 1$. Describe the test statistic and appropriate sampling distribution under the null.

4. In the regression/projection model

$$\begin{aligned} y_i &= x_i'\beta + e_i \\ E\left(x_i e_i\right) &= 0 \end{aligned}$$

the asymptotic distribution for $\hat{\beta}$ is largely determined by that of

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i e_i.$$

Briefly, what is the asymptotic distribution of S_n? (You do not need to re-derive it.)

Now draw n iid $N(0,1)$ random variables u_i, $i = 1, .., n$, independent of the sample. Define

$$S_n^* = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i e_i u_i.$$

What is the exact distribution of S_n^*, conditional on the sample? As $n \to \infty$, what is its asymptotic distribution?