Econometrics, Michaelmas Term 2015, Quiz 3

Exercise 1. Write down the general form of the (conditional) variance-covariance matrix of the error terms. Explain the difference between homoscedastic and heteroscedastic errors.

$$\mathrm{E}\left[UU'|X\right] = \mathrm{E}\left[\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}\begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix}|X\right] = \begin{pmatrix} \mathrm{E}\left[u_1^2|X\right] & \mathrm{E}\left[u_1u_2|X\right] & \dots & \mathrm{E}\left[u_1u_n|X\right] \\ \mathrm{E}\left[u_2u_1|X\right] & \mathrm{E}\left[u_n^2|X\right] & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}\left[u_nu_1|X\right] & \dots & \dots & \mathrm{E}\left[u_n^2|X\right] \end{pmatrix} = \sigma_U^2\Omega$$

Homoscedastic errors: $\sigma_U^2\Omega = \sigma_u^2 I_n$, where $\sigma_u^2 = \mathrm{Var}\left[u_i|X\right]$. Heteroscedastic errors: $\sigma_{u_i}^2 = \mathrm{Var}\left[u_i|X\right] \neq \mathrm{Var}\left[u_j|X\right] = \sigma_{u_j}^2$ for some $i \neq j$.

Exercise 2. Suppose that in a linear regression model the errors are heteroscedastic and that $\mathrm{E}\left[UU'|X\right] = \sigma_U^2\Omega$, where Ω is a positive definite matrix. Derive the expression for the (conditional) variance of the OLS estimator $\hat{\boldsymbol{\beta}}$. Compare this with the (conditional) variance of $\hat{\boldsymbol{\beta}}$, when the errors are homoscedastic. Explain what (if anything) happens to statistical inference based on the usual OLS estimator of $\mathrm{Var}\left(\hat{\boldsymbol{\beta}}|X\right) = \hat{\sigma}_u^2\left(X'X\right)^{-1}$ and why? If something goes wrong, how can it be corrected to provide "good" inference?

$$\begin{aligned} \mathrm{Var}\left(\hat{\boldsymbol{\beta}}|X\right) &= \mathrm{E}\left(\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'|X\right) \\ &= \mathrm{E}\left(\left(X'X\right)^{-1}X'UU'X\left(X'X\right)^{-1}|X\right) \\ &= \left(X'X\right)^{-1}X'\left(\sigma_U^2\Omega\right)X\left(X'X\right)^{-1} \end{aligned}$$

The conditional variance estimator $\hat{\sigma}_u^2\left(X'X\right)^{-1}$ will be biased and so statistical inference based on it may be misleading. Use the Generalized Least Squares and the robust standard errors (HAC).

Exercise 3. Suppose that in a linear regression $y = \beta_0 + \beta_1 x_1^* + u$, the explanatory variable is measured with error and instead $x_1 = x_1^* + v$ is observed, where v is the measurement error s.t. $\mathrm{Cov}\left(x_1^*, v\right) = 0$. Is the OLS estimator $\hat{\beta}_1$ consistent for β_1.

$$y = \beta_1 + \beta_2 x_2 + \left(u - \beta_2 v\right)$$

where $\mathrm{E}\left[u-\beta_1 v\right] = 0$ but $\mathrm{Cov}\left(x_1, u-\beta_1 v\right) = -\beta_1\mathrm{Cov}\left(x_1, v\right) = -\beta_1\mathrm{Var}\left(v\right)$. This implies endogeneity: the error term is correlated with the explanatory variable. Now, we can show that $\hat{\beta}_1$ is inconsistent:

$$p\lim\left(\hat{\beta}_1\right) = \beta_1 + \frac{\mathrm{Cov}\left(x_1, u-\beta_1 v\right)}{\mathrm{Var}\left(x_1\right)} = \beta_1 - \frac{\beta_1\mathrm{Cov}\left(x_1, v\right)}{\mathrm{Var}\left(x_1\right)} = \beta_1\left(\frac{\mathrm{Var}\left(x_1^*\right)}{\mathrm{Var}\left(x_1^*\right)+\mathrm{Var}\left(v\right)}\right)$$

Notice that $0 < \frac{\mathrm{Var}\left(x_1^*\right)}{\mathrm{Var}\left(x_1^*\right)+\mathrm{Var}(v)} < 1$ and so $p\lim\left(\hat{\beta}_1\right)$ is always closer to zero than β_1. This is called the attenuation bias.

Exercise 4. Explain (using the simplest case) the approach of Instrumental Variables for correcting the endogeneity problem. What conditions ensure the IV is "good"?
Consider a simple regression $y = \beta x + u$ where $\mathrm{Cov}\left(x, u\right) \neq 0$, and the instrument $x = \alpha z + e$, which is valid ($\mathrm{Cov}\left(z, u\right) = 0$) and relevant ($\alpha \neq 0$). Then the 2SLS provides an unbiased, consistent and efficient estimate of β.