Economics 522A, Spring 2007

Lecture Note 14: Heteroskedasticity

1 Alternative Variance Matrix: Finite Sample Theory

Suppose

$$V[y|X] = \Omega \neq \sigma^2 I_n.$$

For example, we could have

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & & 0 \\ \vdots & & \ddots & \\ 0 & & & \sigma_n^2 \end{bmatrix}.$$

This says that the variance differs for different units in the sample.

We might also have nonzero terms in the off-diagonal elements of Ω. This would mean that there is correlation across different units in the sample.

If we continue to assume that $E[y|X] = X\beta$, then some of the earlier conclusions about the OLS estimator $\hat{\beta}$ continue to hold, but others do not. (We continue to assume that $(X'X)$ is invertible.)

<u>OLS is unbiased</u>:

$$\begin{aligned} E[\hat{\beta}|X] &= E[(X'X)^{-1}X'y|X] \\ &= (X'X)^{-1}X'E[y|X] \\ &= \beta. \end{aligned}$$

<u>OLS variance changes</u>:

$$\begin{aligned} V[\hat{\beta}|X] &= V[(X'X)^{-1}X'y|X] \\ &= (X'X)^{-1}X'V[y|X]X(X'X)^{-1} \\ &= (X'X)^{-1}X'\Omega X(X'X)^{-1} \end{aligned}$$

<u>Under normality, OLS is still conditionally normal, but with different variance</u>:

If $y|X \sim N(X\beta, \Omega)$, then

$$\hat{\beta}|X \sim N(\beta, (X'X)^{-1}X'\Omega X(X'X)^{-1}).$$

2 Asymptotic Theory for OLS Under Heteroskedasticity

In order to develop large-sample theory, let us continue to assume that $(y_i, x_i')'$ is an IID sample from some joint distribution. We assume that

$$E[y_i|x_i] = x_i'\beta,$$

however, we do not assume that the variance is constant. So there is no single σ^2 such that $V[y_i|x_i] = \sigma^2$. Instead, $V[y_i|x_i]$ could be different depending on the value of x_i. We call this conditional heteroskedasticity.

For example, suppose that

$$V[y_i|x_i] = g(x_i),$$

for some function g. Then the variance matrix of the vector y given X will be

$$V[y|X] = \Omega = \begin{bmatrix} g(x_1) & 0 & \cdots & 0 \\ 0 & g(x_2) & & \\ \vdots & & \ddots & \\ 0 & & & g(x_n) \end{bmatrix}.$$

OLS is consistent:

We continue to assume that the fourth moments of the joint distribution of (y_i, x_i') exist, and that $E[x_i x_i']$ is nonsingular. Then

$$\hat{\beta} = \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' \right)^{-1} \frac{1}{n} \sum_{i=1}^{n} x_i y_i \xrightarrow{\text{p}} E[x_i x_i']^{-1} E[x_i y_i]$$

But, as before,

$$E[x_i y_i] = E[E[x_i y_i | x_i]] = E[x_i E[y_i|x_i]] = E[x_i x_i']\beta.$$

So

$$\hat{\beta} \xrightarrow{\text{p}} \beta.$$

OLS is asymptotically normal, but with different variance:

Let $\varepsilon_i = y_i - x_i'\beta$. Then by our earlier arguments, we can write

$$\sqrt{n}(\hat{\beta} - \beta) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i x_i' \right)^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \varepsilon_i.$$

By the central limit theorem,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \varepsilon_i \xrightarrow{\text{d}} N(0, V(x_i \varepsilon_i)),$$

where

$$V(x_i\varepsilon_i) = E[\varepsilon_i^2 x_i x_i'],$$

since $E[x_i\varepsilon_i] = 0$. However, we cannot simplify the variance any further than this. So, by Slutsky's lemma,

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\text{d}} N(0, W),$$

where

$$W = E[x_i x_i']^{-1} E[\varepsilon_i^2 x_i x_i'] E[x_i x_i']^{-1}.$$

Since the limiting variance is no longer $\sigma^2 E[x_i x_i']^{-1}$, the usual standard errors we calculate for the OLS estimator are no longer correct, and the usual test statistics and confidence intervals are not valid.

The "bias" of the usual standard errors can be either positive or negative, depending on the specific sample, and it is difficult to know exactly how our inference will be affected if we use the wrong variance estimates.

Fortunately, we can construct an alternative variance estimate that is consistent under heteroskedasticity. The idea is to simply replace expectations by sample averages:

$$\widehat{W} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} e_i^2 x_i x_i'\right) \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1},$$

where $e_i = y_i - x_i'\hat{\beta}$ are the OLS residuals.

This is called the Eicker-White variance estimator. We can use this to construct standard errors for the OLS estimator that are "robust" against possible heteroskedasticity. So we sometimes speak of "robust standard errors" or "heteroskedascity-consistent standard errors."

Since many economic data sets appear to exhibit heteroskedasticity, it is often a good idea to calculate the robust standard errors and report these in addition to (or in place of) the conventional standard errors that assume $V[y_i|x_i] = \sigma^2$.

The Eicker-White variance estimator can be somewhat biased in small to moderate samples. There is a small correction that is often applied; see Ruud p.429 for more details.