16.3 Covariance Matrix Estimation

At first glance, the generalized OLS covariance matrix estimator and its NLS analog (16.08) do not seem to be very useful. To compute them we need to know $\boldsymbol{\Omega}$, but if we knew $\boldsymbol{\Omega}$, we could use GLS or GNLS and obtain more efficient estimates. This was the conventional wisdom among econometricians until a decade ago. But an extremely influential paper by White (1980) showed that it is in fact possible to obtain an estimator of the covariance matrix of least squares estimates that is asymptotically valid when there is heteroskedasticity of unknown form.[2] Such an estimator is a called a **heteroskedasticity-consistent covariance matrix estimator**, or **HCCME**.

The key to obtaining an HCCME is to recognize that we do *not* have to estimate $\boldsymbol{\Omega}$ consistently. That would indeed be an impossible task, since $\boldsymbol{\Omega}$ has n diagonal elements to estimate. The asymptotic covariance matrix of a vector of NLS estimates, under heteroskedasticity, is given by expression (16.08), which can be rewritten as

$$\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\boldsymbol{X}_0^{\top}\boldsymbol{X}_0\right)^{-1}\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\boldsymbol{X}_0^{\top}\boldsymbol{\Omega}\boldsymbol{X}_0\right)\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\boldsymbol{X}_0^{\top}\boldsymbol{X}_0\right)^{-1}. \tag{16.09}$$

The first and third factors here are identical, and we can easily estimate them in the usual way. A consistent estimator is

$$\frac{1}{n}\hat{\boldsymbol{X}}^{\top}\hat{\boldsymbol{X}},$$

where $\hat{\boldsymbol{X}} \equiv \boldsymbol{X}(\hat{\boldsymbol{\beta}})$. The only tricky thing, then, is to estimate the second factor. White showed that this second factor can be estimated consistently by

$$\frac{1}{n}\hat{\boldsymbol{X}}^{\top}\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{X}}, \tag{16.10}$$

where $\hat{\boldsymbol{\Omega}}$ may be any of several different *inconsistent* estimators of $\boldsymbol{\Omega}$. The simplest version of $\hat{\boldsymbol{\Omega}}$, and the one that White proposed in the context of linear regression models, has t^{th} diagonal element equal to $\hat{u}_t^2$, the t^{th} squared least squares residual.

Unlike $\boldsymbol{\Omega}$, the middle factor of (16.09) has only $\frac{1}{2}(k^2+k)$ distinct elements, whatever the sample size. That is why it is possible to estimate it consistently. A typical element of this matrix is

$$\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\sum_{t=1}^{n}\omega_t^2 X_{ti}X_{tj}\right), \tag{16.11}$$

[2] Precursors of White's paper in the statistics literature include Eicker (1963, 1967) and Hinkley (1977), as well as some of the early papers on bootstrapping (see Chapter 21).