This equation defines an ordinary IV estimator in terms of the transformed variables $\boldsymbol{y}^*$ and $\boldsymbol{X}^*$ and the transformed instruments $\boldsymbol{Z}$. Thus the estimator defined by (17.53) can be calculated with no more difficulty than the GLS estimator. It is appropriate to use it when GLS or feasible GLS would have been appropriate except for possible correlation of the error terms with the regressors.

The estimator defined by (17.53) bears a close resemblance to the H2SLS estimator (17.44) defined in the last section. In fact, replacing $\boldsymbol{W}$ in the latter by $\boldsymbol{\Omega}^{-1}\boldsymbol{W}$ yields the former. The theory developed in this section shows that if it is possible to choose $\boldsymbol{W}$ as the conditional expectations of the regressors $\boldsymbol{X}$ (or linear combinations of them), then the estimator defined by (17.53) is asymptotically efficient, and the H2SLS estimator is not. The advantage of H2SLS is that it can be calculated in the presence of heteroskedasticity of unknown form, since $n^{-1}\boldsymbol{W}^{\top}\boldsymbol{\Omega}\boldsymbol{W}$ can be estimated consistently by use of inconsistent estimators of $\boldsymbol{\Omega}$. (17.53), on the other hand, can be formulated only if $\boldsymbol{\Omega}$ itself can be consistently estimated, because expressions like $n^{-1}\boldsymbol{W}^{\top}\boldsymbol{\Omega}^{-1}\boldsymbol{W}$ and $n^{-1}\boldsymbol{W}^{\top}\boldsymbol{\Omega}^{-1}\boldsymbol{y}$ *cannot* be estimated consistently without a consistent estimate of $\boldsymbol{\Omega}$. Thus both estimators are useful, but in different circumstances.

The concept of the GMM bound was introduced, not under that name, by Hansen (1985), who also provided conditions for optimal instruments. The arguments used in order to derive the bound have a longer history, however, and Hansen traces the history of the search for efficient instruments back as far as Basmann (1957) and Sargan (1958).

17.5 Covariance Matrix Estimation

In previous sections, we mentioned the difficulties that can arise in estimating covariance matrices in the GMM context. In fact, problems occur at two distinct points: once for the choice of the weighting matrix to be used in constructing a criterion function and again for estimating the asymptotic covariance matrix of the estimates. Fortunately, similar considerations apply to both problems, and so we can consider them together.

Recall from (17.31) that the asymptotic covariance matrix of a GMM estimator computed using a weighting matrix $\boldsymbol{A}_0$ is

$$\big(\boldsymbol{D}^{\top}\boldsymbol{A}_0\boldsymbol{D}\big)^{-1}\boldsymbol{D}^{\top}\boldsymbol{A}_0\boldsymbol{\Phi}\boldsymbol{A}_0\boldsymbol{D}\big(\boldsymbol{D}^{\top}\boldsymbol{A}_0\boldsymbol{D}\big)^{-1},$$

in the notation of Section 17.2. If the necessary condition for efficiency of Theorem 17.3 is to be satisfied, it is required that $\boldsymbol{A}_0 \overset{a}{=} \boldsymbol{\Phi}^{-1}$, where $\boldsymbol{\Phi}$ is the $l \times l$ asymptotic covariance matrix of the empirical moments $n^{-1/2}\boldsymbol{F}^{\top}(\boldsymbol{\theta})\boldsymbol{\iota}$ with typical element

$$n^{-1/2}\sum_{t=1}^{n} f_{ti}(y_t, \boldsymbol{\theta}).$$