squares. The first property is that, if (8.49) is evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_{\rm IV}$, then the regressors $\boldsymbol{P_W X}$ are orthogonal to the regressand, because the orthogonality conditions, namely,

$$\boldsymbol{X}^\top \boldsymbol{P_W}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\rm IV}) = \boldsymbol{0},$$

are just the moment conditions (8.28) that define $\hat{\boldsymbol{\beta}}_{\rm IV}$.

The second property is that, if (8.49) is again evaluated at $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_{\rm IV}$, the estimated OLS covariance matrix is asymptotically valid. This matrix is

$$s^2(\boldsymbol{X}^\top \boldsymbol{P_W X})^{-1}. \tag{8.50}$$

Here s^2 is the sum of squared residuals from (8.49), divided by $n - k$. Since $\hat{\boldsymbol{b}} = \boldsymbol{0}$ because of the orthogonality of the regressand and the regressors, those residuals are the components of the vector $\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\rm IV}$, that is, the IV residuals from (8.10). It follows that (8.50), which has exactly the same form as (8.34), is a consistent estimator of the covariance matrix of $\hat{\boldsymbol{\beta}}_{\rm IV}$, where "consistent estimator" is used in the sense of (5.22). As with the ordinary GNR, the estimator $\acute{s}^2$ obtained by running (8.49) with $\boldsymbol{\beta} = \acute{\boldsymbol{\beta}}$ is consistent for the error variance σ^2 if $\acute{\boldsymbol{\beta}}$ is root-n consistent; see Exercise 8.16.

The third property is that, like the ordinary GNR, the IVGNR permits one-step efficient estimation. For linear models, this is true if *any* value of $\boldsymbol{\beta}$ is used in (8.49). If we set $\boldsymbol{\beta} = \acute{\boldsymbol{\beta}}$, then running (8.49) gives the artificial parameter estimates

$$\acute{\boldsymbol{b}} = (\boldsymbol{X}^\top \boldsymbol{P_W X})^{-1}\boldsymbol{X}^\top \boldsymbol{P_W}(\boldsymbol{y} - \boldsymbol{X}\acute{\boldsymbol{\beta}}) = \hat{\boldsymbol{\beta}}_{\rm IV} - \acute{\boldsymbol{\beta}},$$

from which it follows that $\acute{\boldsymbol{\beta}} + \acute{\boldsymbol{b}} = \hat{\boldsymbol{\beta}}_{\rm IV}$ for all $\acute{\boldsymbol{\beta}}$. In the context of nonlinear IV estimation (see Section 8.9), this result, like the one above for $\acute{s}^2$, becomes an approximation that is asymptotically valid only if $\acute{\boldsymbol{\beta}}$ is a root-n consistent estimator of the true $\boldsymbol{\beta}_0$.

Tests Based on the IVGNR

If the restrictions to be tested are all linear restrictions, there is no further loss of generality if we suppose that they are all zero restrictions. Thus the null and alternative hypotheses can be written as

$$H_0\colon\ \boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{u}, \text{ and} \tag{8.51}$$

$$H_1\colon\ \boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u}, \tag{8.52}$$

where the matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are, respectively, $n \times k_1$ and $n \times k_2$, $\boldsymbol{\beta}_1$ is a k_1–vector, and $\boldsymbol{\beta}_2$ is a k_2–vector. As elsewhere in this chapter, it is assumed that $\mathrm{E}(\boldsymbol{uu}^\top) = \sigma^2\mathbf{I}$. Any or all of the columns of $\boldsymbol{X} = [\boldsymbol{X}_1 \;\; \boldsymbol{X}_2]$ may be correlated with the error terms. It is assumed that there exists an $n \times l$ matrix $\boldsymbol{W}$ of instruments, which are asymptotically uncorrelated with the error terms, and that $l \geq k = k_1 + k_2$.