Solution to Exercise 8.24

★**8.24** Using the same methods as those in Sections 6.5 and 6.6, show that the nonlinear version (8.89) of the IVGNR satisfies the three conditions, analogous to those set out in Exercise 8.20, which are necessary for the use of the IVGNR in hypothesis testing. What is the nonlinear version of the IV variant of the HRGNR? Show that it, too, satisfies the three conditions under the assumption of possibly heteroskedastic error terms.

The nonlinear version of the IVGNR is

$$\boldsymbol{y} - \boldsymbol{x}(\boldsymbol{\beta}) = \boldsymbol{P_W X}(\boldsymbol{\beta})\boldsymbol{b} + \text{residuals}. \tag{8.89}$$

When it is evaluated at $\hat{\boldsymbol{\beta}}_{\rm IV}$, the inner product of the regressand with the matrix of regressors is

$$(\boldsymbol{y} - \hat{\boldsymbol{x}})^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}},$$

where $\hat{\boldsymbol{x}} \equiv \boldsymbol{x}(\hat{\boldsymbol{\beta}}_{\rm IV})$ and $\hat{\boldsymbol{X}} \equiv \boldsymbol{X}(\hat{\boldsymbol{\beta}}_{\rm IV})$. This inner product is equal to $\mathbf{0}$ because $\hat{\boldsymbol{\beta}}_{\rm IV}$ must satisfy the moment conditions (8.84). Thus the IVGNR satisfies the first condition necessary for its use in hypothesis testing.

Since the regressors of the IVGNR evaluated at $\hat{\boldsymbol{\beta}}_{\rm IV}$ have no explanatory power for the regressand, the OLS covariance matrix estimator is

$$\frac{n}{n-k}\,\hat{\sigma}^2(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1}.$$

This is $n/(n-k)$ times the covariance matrix estimator (8.88). The factor of $n/(n-k)$ arises because the OLS estimator of σ^2 divides by $n-k$, and the IV estimator divided by n. Thus the IVGNR also satisfies the second condition.

For the third condition, we need to show that, if the IVGNR is evaluated at a root-n consistent estimator $\acute{\boldsymbol{\beta}}$, then

$$n^{1/2}(\hat{\boldsymbol{\beta}}_{\rm IV} - \boldsymbol{\beta}_0) \overset{a}{=} n^{1/2}(\acute{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + n^{1/2}\acute{\boldsymbol{b}}. \tag{S8.34}$$

By the standard formula for the OLS estimator,

$$\begin{aligned}\acute{\boldsymbol{b}} &= (\acute{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\acute{\boldsymbol{X}})^{-1}\acute{\boldsymbol{X}}^{\top}\boldsymbol{P_W}(\boldsymbol{y} - \acute{\boldsymbol{x}})\\ &= (\acute{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\acute{\boldsymbol{X}})^{-1}\acute{\boldsymbol{X}}^{\top}\boldsymbol{P_W}(\boldsymbol{x}_0 - \acute{\boldsymbol{x}} + \boldsymbol{u}).\end{aligned} \tag{S8.35}$$

Since $\acute{\boldsymbol{\beta}}$ is consistent, a first-order Taylor series approximation shows that

$$n^{1/2}(\acute{\boldsymbol{x}} - \boldsymbol{x}_0) \overset{a}{=} n^{1/2}\boldsymbol{X}_0(\acute{\boldsymbol{\beta}} - \boldsymbol{\beta}_0).$$

Substituting this into (S8.35) and adding appropriate powers of n yields the result that

$$n^{1/2}\acute{\boldsymbol{b}} \stackrel{a}{=} (n^{-1}\acute{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\acute{\boldsymbol{X}})^{-1}n^{-1}\acute{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\big(n^{1/2}\boldsymbol{X}_0(\boldsymbol{\beta}_0 - \acute{\boldsymbol{\beta}}) + n^{1/2}\boldsymbol{u}\big). \tag{S8.36}$$

Because $\acute{\boldsymbol{\beta}}$ is assumed to be root-n consistent, we can replace $\acute{\boldsymbol{X}}$ here by $\boldsymbol{X}_0$ without changing the asymptotic equality. Therefore, (S8.36) implies that

$$n^{1/2}\acute{\boldsymbol{b}} \stackrel{a}{=} (n^{-1}\boldsymbol{X}_0{}^{\top}\boldsymbol{P_W}\boldsymbol{X}_0)^{-1}n^{-1/2}\boldsymbol{X}_0{}^{\top}\boldsymbol{P_W}\boldsymbol{u} - n^{1/2}(\acute{\boldsymbol{\beta}} - \boldsymbol{\beta}_0). \tag{S8.37}$$

Here we have used the fact that $(n^{-1}\boldsymbol{X}_0{}^{\top}\boldsymbol{P_W}\boldsymbol{X}_0)^{-1}(n^{-1}\boldsymbol{X}_0{}^{\top}\boldsymbol{P_W}\boldsymbol{X}_0) = \mathbf{I}$. As was noted in (8.86), the first term on the right-hand side of (S8.37) is asymptotically equal to $n^{1/2}(\hat{\boldsymbol{\beta}}_{\rm IV} - \boldsymbol{\beta}_0)$. Therefore, (S8.37) implies that

$$n^{1/2}\acute{\boldsymbol{b}} \stackrel{a}{=} n^{1/2}(\hat{\boldsymbol{\beta}}_{\rm IV} - \boldsymbol{\beta}_0) - n^{1/2}(\acute{\boldsymbol{\beta}} - \boldsymbol{\beta}_0),$$

which can be rearranged to yield (S8.34).

The nonlinear version of the IV variant of the HRGNR is

$$\boldsymbol{\iota} = \boldsymbol{P}_{\boldsymbol{U}(\boldsymbol{\beta})\boldsymbol{P_W}\boldsymbol{X}(\boldsymbol{\beta})}\boldsymbol{U}^{-1}(\boldsymbol{\beta})\boldsymbol{P_W}\boldsymbol{X}(\boldsymbol{\beta})\boldsymbol{b} + \text{residuals}, \tag{S8.38}$$

where $\boldsymbol{U}(\boldsymbol{\beta})$ is an $n \times n$ diagonal matrix, the $t^{\rm th}$ diagonal element of which is $y_t - x_t(\boldsymbol{\beta})$.

When (S8.38) is evaluated at $\hat{\boldsymbol{\beta}}_{\rm IV}$, the inner product of the regressand with the matrix of regressors is

$$\begin{aligned}
&\boldsymbol{\iota}^{\top}\boldsymbol{P}_{\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}}}\hat{\boldsymbol{U}}^{-1}\boldsymbol{P_W}\hat{\boldsymbol{X}}\\
&\quad = \boldsymbol{\iota}^{\top}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1}\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}^{-1}\boldsymbol{P_W}\hat{\boldsymbol{X}}\\
&\quad = \hat{\boldsymbol{u}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1}\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}}.
\end{aligned} \tag{S8.39}$$

Since the moment conditions (8.84) imply that $\hat{\boldsymbol{u}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}} = \mathbf{0}$, the last line of (S8.39) is just a zero vector. Therefore, as required, we have shown that the regressand in (S8.38) is orthogonal to the regressors when $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\rm IV}$.

The estimated OLS covariance matrix from (8.90), evaluated at $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\rm IV}$, is

$$\frac{n}{n-k}\big(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}^{-1}\boldsymbol{P}_{\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}}}\hat{\boldsymbol{U}}^{-1}\boldsymbol{P_W}\hat{\boldsymbol{X}}\big)^{-1}.$$

The first factor here is $\boldsymbol{\iota}^{\top}\boldsymbol{\iota}/(n-k)$, which is the OLS estimate of σ^2 from regression (S8.38). Because the regressand is orthogonal to the regressors, the SSR is precisely $\boldsymbol{\iota}^{\top}\boldsymbol{\iota}$. The second factor can be rewritten as

$$\begin{aligned}
&\big(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}^{-1}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1}\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}^{-1}\boldsymbol{P_W}\hat{\boldsymbol{X}}\big)^{-1}\\
&\qquad = \big(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1}\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}}\big)^{-1}\\
&\qquad = (\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1}\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{\Omega}}\boldsymbol{P_W}\hat{\boldsymbol{X}}(\hat{\boldsymbol{X}}^{\top}\boldsymbol{P_W}\hat{\boldsymbol{X}})^{-1},
\end{aligned}$$

which is the analog, for nonlinear IV estimation, of the HCCME $\widehat{\mathrm{Var}}_{\mathrm{h}}(\hat{\boldsymbol{\beta}}_{\mathrm{IV}})$ given by (8.65).

Finally, we need to show that (S8.38) satisfies the one-step property. By exactly the same reasoning as we used in Exercise 8.20 to prove the analogous result for the HRGNR in the linear IV case, it can easily be shown that

$$\acute{\boldsymbol{b}} = (\acute{\boldsymbol{X}}^{\top}\boldsymbol{P}_{\boldsymbol{W}}\acute{\boldsymbol{X}})^{-1}\acute{\boldsymbol{X}}^{\top}\boldsymbol{P}_{\boldsymbol{W}}(\boldsymbol{x}_0 - \acute{\boldsymbol{x}} + \boldsymbol{u}).$$

This is precisely equation (S8.35). The rest of the argument for the nonlinear IVGNR (8.89) goes through unchanged in the heteroskedasticity-robust case, and we conclude that (S8.34) holds for the heteroskedasticity-robust IVGNR (S8.38) just as it does for the ordinary IVGNR (8.89).