Solution to Exercise 8.20

$\star$**8.20** The IV variant of the HRGNR (6.90), evaluated at $\boldsymbol{\beta} = \acute{\boldsymbol{\beta}}$, can be written as

$$\boldsymbol{\iota} = \boldsymbol{P}_{\acute{\boldsymbol{U}}\boldsymbol{P_W X}} \acute{\boldsymbol{U}}^{-1} \boldsymbol{P_W X b} + \text{residuals}, \tag{8.90}$$

where $\boldsymbol{\iota}$ is an n–vector of which every component equals 1, and $\acute{\boldsymbol{U}}$ is an $n \times n$ diagonal matrix with t^{th} diagonal element equal to the t^{th} element of the vector $\boldsymbol{y} - \boldsymbol{X}\acute{\boldsymbol{\beta}}$.

Verify that this artificial regression possesses all the requisite properties for hypothesis testing, namely, that

- The regressand in (8.90) is orthogonal to the regressors when $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{IV}}$;
- The estimated OLS covariance matrix from (8.90) evaluated at $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{IV}}$ is equal to $n/(n-k)$ times the HCCME $\widehat{\text{Var}}_{\text{h}}(\hat{\boldsymbol{\beta}}_{\text{IV}})$ given by (8.65);
- The HRGNR (8.90) allows one-step estimation: The OLS parameter estimates $\acute{\boldsymbol{b}}$ from (8.90) are such that $\hat{\boldsymbol{\beta}}_{\text{IV}} = \acute{\boldsymbol{\beta}} + \acute{\boldsymbol{b}}$.

The first result is easily shown. When $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{IV}}$, the inner product of the regressand of (8.90) with the matrix of regressors is

$$\begin{aligned} &\boldsymbol{\iota}^\top \boldsymbol{P}_{\hat{\boldsymbol{U}}\boldsymbol{P_W X}} \hat{\boldsymbol{U}}^{-1} \boldsymbol{P_W X} \\ &\quad = \boldsymbol{\iota}^\top \hat{\boldsymbol{U}} \boldsymbol{P_W X} (\boldsymbol{X}^\top \boldsymbol{P_W} \hat{\boldsymbol{U}}\hat{\boldsymbol{U}} \boldsymbol{P_W X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \hat{\boldsymbol{U}}\hat{\boldsymbol{U}}^{-1} \boldsymbol{P_W X} \\ &\quad = \hat{\boldsymbol{u}}^\top \boldsymbol{P_W X} (\boldsymbol{X}^\top \boldsymbol{P_W} \hat{\boldsymbol{U}}\hat{\boldsymbol{U}} \boldsymbol{P_W X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W X}. \end{aligned} \tag{S8.25}$$

The matrix $\hat{\boldsymbol{U}}$ is not explicitly transposed, because it is a diagonal matrix. In the last line of (S8.25), $\hat{\boldsymbol{u}} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{\text{IV}}$. Since the moment conditions (8.28) imply that $\hat{\boldsymbol{u}}^\top \boldsymbol{P_W X} = \boldsymbol{0}$, this last line is just a zero vector. Therefore, as required, we have shown that the regressand in (8.90) is orthogonal to the regressors when $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{IV}}$.

The second result is also easily shown. The OLS covariance matrix estimate from (8.90), evaluated at $\acute{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{IV}}$, is

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

The first factor here is $\boldsymbol{\iota}^\top\boldsymbol{\iota}/(n-k)$, which is the OLS estimate of σ^2 from regression (8.90). Because the regressand is orthogonal to the regressors, the SSR is precisely $\boldsymbol{\iota}^\top\boldsymbol{\iota}$. The second factor can be rewritten as

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

which is the HCCME $\widehat{\text{Var}}_{\text{h}}(\hat{\boldsymbol{\beta}}_{\text{IV}})$ given by (8.65).

For the third result, we need to show that $\hat{\boldsymbol{\beta}}_{\rm IV} = \acute{\boldsymbol{\beta}} + \acute{\boldsymbol{b}}$. By the standard formula for the OLS estimator,

$$\acute{\boldsymbol{b}} = \big(\boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{U}}^{-1} \boldsymbol{P}_{\acute{\boldsymbol{U}}\boldsymbol{P_W}\boldsymbol{X}} \acute{\boldsymbol{U}}^{-1} \boldsymbol{P_W} \boldsymbol{X}\big)^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{U}}^{-1} \boldsymbol{P}_{\acute{\boldsymbol{U}}\boldsymbol{P_W}\boldsymbol{X}} \boldsymbol{\iota}.$$

We have just seen that the first factor here can be rewritten as

$$(\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{\Omega}} \boldsymbol{P_W} \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1}. \tag{S8.26}$$

The second factor is

$$\begin{aligned} &\boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{U}}^{-1} \acute{\boldsymbol{U}} \boldsymbol{P_W} \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{U}} \acute{\boldsymbol{U}} \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{U}} \boldsymbol{\iota} \\ &\quad = \boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{\Omega}} \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{u}}. \end{aligned} \tag{S8.27}$$

Postmultiplying (S8.26) by (S8.27), we observe that the last factor in (S8.26) is the inverse of the first factor in (S8.27), so that

$$\begin{aligned} \acute{\boldsymbol{b}} &= (\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{\Omega}} \boldsymbol{P_W} \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{\Omega}} \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{u}} \\ &= (\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \acute{\boldsymbol{u}}. \end{aligned}$$

It is then easy to see that

$$\begin{aligned} \acute{\boldsymbol{b}} &= (\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} (\boldsymbol{y} - \boldsymbol{X} \acute{\boldsymbol{\beta}}) \\ &= (\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{y} - (\boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{P_W} \boldsymbol{X} \acute{\boldsymbol{\beta}} \\ &= \hat{\boldsymbol{\beta}}_{\rm IV} - \acute{\boldsymbol{\beta}}. \end{aligned}$$

This establishes the third result