Solution to Exercise 9.18

$\star$**9.18** Derive a heteroskedasticity robust version of the artificial regression (9.125), assuming that the covariance matrix of the vector $\boldsymbol{f}(\boldsymbol{\theta})$ of zero functions is diagonal, but otherwise arbitrary.

The heteroskedasticity robust version of the artificial regression (9.125) looks very much like the HRGNR for IV linear regression; see Exercise 8.20. It can be written as

$$\boldsymbol{\iota} = -\boldsymbol{P}_{\boldsymbol{U}(\boldsymbol{\theta})\boldsymbol{P_W}\boldsymbol{F}(\boldsymbol{\theta})}\boldsymbol{U}^{-1}(\boldsymbol{\theta})\boldsymbol{P_W}\boldsymbol{F}(\boldsymbol{\theta})\boldsymbol{b} + \text{residuals}, \tag{S9.42}$$

where $\boldsymbol{U}(\boldsymbol{\theta})$ is a diagonal matrix with typical diagonal element $f_t(\boldsymbol{\theta})$.

If we evaluate (S9.42) at the GMM estimates $\hat{\boldsymbol{\theta}}$ which minimize the criterion function (9.124), we find that the transpose of the regressand times the matrix of regressors is

$$\begin{aligned}
&\boldsymbol{\iota}^\top\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{F}}(\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{F}})^{-1}\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}^{-1}\hat{\boldsymbol{F}}\\
&\quad = \hat{\boldsymbol{f}}^\top\boldsymbol{P_W}\hat{\boldsymbol{F}}(\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{U}}\hat{\boldsymbol{U}}\boldsymbol{P_W}\hat{\boldsymbol{F}})^{-1}\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{F}}.
\end{aligned}$$

It is evident that the regressand will be orthogonal to the regressors whenever the conditions

$$\boldsymbol{F}^\top(\hat{\boldsymbol{\theta}})\boldsymbol{P_W}\boldsymbol{f}(\hat{\boldsymbol{\theta}}) = \boldsymbol{0}$$

are satisfied. These are just the first-order conditions that correspond to the criterion function (9.124).

It is also not difficult to see that the OLS covariance matrix from regression (S9.42) evaluated at $\hat{\boldsymbol{\theta}}$ is

$$\frac{n}{n-k}(\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{F}})^{-1}\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{\Omega}}\boldsymbol{P_W}\hat{\boldsymbol{F}}(\hat{\boldsymbol{F}}^\top\boldsymbol{P_W}\hat{\boldsymbol{F}})^{-1},$$

where $\hat{\boldsymbol{\Omega}} \equiv \hat{\boldsymbol{U}}\hat{\boldsymbol{U}}$; compare (8.65). In addition, one-step estimation can be shown to work by adding asymptotic details to the argument used in Section 6.8 for the HRGNR.