by the ordinary IV procedure. Next, the residuals $\hat{u}_t$ are combined with the instruments in the direct product $\hat{\boldsymbol{V}} \equiv \hat{\boldsymbol{u}} * \boldsymbol{W}$. Then $\hat{\boldsymbol{\Gamma}}(j)$ is n^{-1} times the $l \times l$ matrix of inner products of the columns of $\hat{\boldsymbol{V}}$ with these same columns lagged j times, the initial unobserved elements being replaced by zeros. As we saw above, $\hat{\boldsymbol{\Gamma}}(0)$ is just $n^{-1}\boldsymbol{W}^{\top}\hat{\boldsymbol{\Omega}}\boldsymbol{W}$, where $\hat{\boldsymbol{\Omega}} = \mathrm{diag}(\hat{u}_t^2)$. Finally, $\hat{\boldsymbol{\Phi}}$ is formed by use of (17.64).

As before, the $\hat{\boldsymbol{\Phi}}$ thus obtained can be used for two purposes. One is to form what is called a **heteroskedasticity and autocorrelation consistent**, or **HAC**, covariance matrix estimator for the ordinary IV estimator. Since the IV estimator is based on the empirical moments $\boldsymbol{W}^{\top}(\boldsymbol{y} - \boldsymbol{X\beta})$ and the weighting matrix $(\boldsymbol{W}^{\top}\boldsymbol{W})^{-1}$, as can be seen from (17.09), the HAC covariance matrix estimator is found by applying the formula (17.31) to the present case and using (17.33) and (17.34). We obtain

$$\big(\boldsymbol{X}^{\top}\boldsymbol{P}_W\boldsymbol{X}\big)^{-1}\boldsymbol{X}^{\top}\boldsymbol{W}\big(\boldsymbol{W}^{\top}\boldsymbol{W}\big)^{-1}n\hat{\boldsymbol{\Phi}}\big(\boldsymbol{W}^{\top}\boldsymbol{W}\big)^{-1}\boldsymbol{W}^{\top}\boldsymbol{X}\big(\boldsymbol{X}^{\top}\boldsymbol{P}_W\boldsymbol{X}\big)^{-1}. \quad (17.65)$$

In the simple case in which $\boldsymbol{W} = \boldsymbol{X}$, this rather complicated formula becomes

$$\big(\boldsymbol{X}^{\top}\boldsymbol{X}\big)^{-1}n\hat{\boldsymbol{\Phi}}\big(\boldsymbol{X}^{\top}\boldsymbol{X}\big)^{-1}.$$

When there is no serial correlation, implying that $n\hat{\boldsymbol{\Phi}} = \boldsymbol{W}^{\top}\hat{\boldsymbol{\Omega}}\boldsymbol{W}$, this simplifies to the familiar HCCME (16.15), specialized to the case of a linear regression model. It is a good exercise to see what (17.65) reduces to when there is no serial correlation and $\boldsymbol{W} \neq \boldsymbol{X}$.

More interesting than the HAC covariance matrix estimator is the estimator analogous to the H2SLS estimator, (17.44). For this, instead of using $(\boldsymbol{W}^{\top}\boldsymbol{W})^{-1}$ as weighting matrix, we use the inverse of $\hat{\boldsymbol{\Phi}}$, calculated in the manner described above by use of the ordinary IV estimator as the preliminary consistent estimator. The criterion function becomes

$$(\boldsymbol{y} - \boldsymbol{X\beta})^{\top}\boldsymbol{W}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{W}^{\top}(\boldsymbol{y} - \boldsymbol{X\beta}),$$

and the estimator, which is sometimes called **two-step two-stage least squares**, is therefore

$$\hat{\boldsymbol{\beta}} = \big(\boldsymbol{X}^{\top}\boldsymbol{W}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{W}^{\top}\boldsymbol{X}\big)^{-1}\boldsymbol{X}^{\top}\boldsymbol{W}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{W}^{\top}\boldsymbol{y}. \quad (17.66)$$

This is very similar to (17.44), in which the matrix $\hat{\boldsymbol{\Phi}}$ is replaced by $\boldsymbol{W}^{\top}\hat{\boldsymbol{\Omega}}\boldsymbol{W}$. Indeed, in the absence of autocorrelation, $n^{-1}\boldsymbol{W}^{\top}\hat{\boldsymbol{\Omega}}\boldsymbol{W}$ is the appropriate estimator of $\boldsymbol{\Phi}$. It is easier to obtain an estimate of the asymptotic covariance matrix of (17.66) than of the ordinary IV estimator. It is simply

$$\hat{\boldsymbol{V}}(\hat{\boldsymbol{\beta}}) = \big(\boldsymbol{X}^{\top}\boldsymbol{W}\hat{\boldsymbol{\Phi}}^{-1}\boldsymbol{W}^{\top}\boldsymbol{X}\big)^{-1}.$$

So far, there is very little practical experience of the estimator (17.66). One reason for this is that econometricians often prefer to model dynamics explicitly (see Chapter 19) rather than leaving all the dynamics in the error term