In order to make (18.61) operational, we need to estimate the covariance matrix $\boldsymbol{\Sigma}$ of the error terms. In the case of an SUR model, we could use OLS on each equation individually. Since OLS is inconsistent for simultaneous equations models, we use 2SLS on each equation instead. Thus the first two "stages" of 3SLS are simply the two stages of 2SLS, applied to each separate equation of (18.01). The covariances of the error terms are then estimated from the 2SLS residuals:

$$\tilde{\sigma}_{ij} = \frac{1}{n}\sum_{t=1}^{n} \tilde{u}_{ti}\tilde{u}_{tj}. \tag{18.62}$$

Of course, these residuals must be the genuine 2SLS residuals, not the residuals from OLS estimation of the second-stage regressions; see Section 7.5. Thus we see that the 3SLS estimators $\tilde{\boldsymbol{\delta}}_1$ through $\tilde{\boldsymbol{\delta}}_g$ must jointly solve the first-order conditions

$$\sum_{j=1}^{g} \tilde{\sigma}^{ij} \boldsymbol{Z}_i^{\top}\boldsymbol{P}_X\big(\boldsymbol{y}_j - \boldsymbol{Z}_j\tilde{\boldsymbol{\delta}}_j\big) = \mathbf{0}. \tag{18.63}$$

The solution is easy to write down. If $\boldsymbol{\delta} \equiv [\boldsymbol{\delta}_1 \vdots \cdots \vdots \boldsymbol{\delta}_g]$ and matrices enclosed in square brackets $[\cdot]$ denote partitioned matrices characterized by a typical block, then the 3SLS estimator $\tilde{\boldsymbol{\delta}}$ can be written very compactly as

$$\tilde{\boldsymbol{\delta}} = \big[\tilde{\sigma}^{ij}\boldsymbol{Z}_i^{\top}\boldsymbol{P}_X\boldsymbol{Z}_j\big]^{-1}\left[\sum_{j=1}^{g}\tilde{\sigma}^{ij}\boldsymbol{Z}_i^{\top}\boldsymbol{P}_X\boldsymbol{y}_j\right]. \tag{18.64}$$

It is more common to see the 3SLS estimator written using an alternative notation that involves Kronecker products; see almost any econometrics textbook. Although Kronecker products can sometimes be useful (Magnus and Neudecker, 1988), we prefer the compact notation of (18.64).

The 3SLS estimator is closely related both to the 2SLS estimator and to the GLS estimator for multivariate SUR models in which the explanatory variables are all exogenous or predetermined. If we assume that $\boldsymbol{\Sigma}$ is diagonal, conditions (18.63) become simply

$$\tilde{\sigma}^{ii}\boldsymbol{Z}_i^{\top}\boldsymbol{P}_X\big(\boldsymbol{y}_i - \boldsymbol{Z}_i\boldsymbol{\delta}_i\big) = \mathbf{0},$$

which are equivalent to the conditions for equation-by-equation 2SLS. Thus 3SLS and 2SLS will be asymptotically (but not numerically) equivalent when the structural form errors are not contemporaneously correlated. It is also easy to see that the SUR estimator for linear models is just a special case of the 3SLS estimator. Since all regressors can be used as instruments in the SUR case, it is no longer necessary to use 2SLS in the preliminary stage. Equivalently, the fact that each regressor matrix $\boldsymbol{Z}_i$ is just a submatrix of the full regressor matrix, $\boldsymbol{X}$, implies that $\boldsymbol{P}_X\boldsymbol{Z}_i = \boldsymbol{Z}_i$. Thus (18.63) simplifies to

$$\sum_{j=1}^{g}\tilde{\sigma}^{ij}\boldsymbol{Z}_i^{\top}\big(\boldsymbol{y}_j - \boldsymbol{Z}_j\boldsymbol{\delta}_j\big) = \mathbf{0},$$