variables in the entire system. Then the second-stage regression for $\boldsymbol{y}$ can simply be written as

$$\boldsymbol{y} = \boldsymbol{P}_W \boldsymbol{X}\boldsymbol{\beta} + \text{residuals}. \tag{7.28}$$

The OLS estimator of $\boldsymbol{\beta}$ from this regression is just the IV estimator (7.17):

$$\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{X}^\top \boldsymbol{P}_W \boldsymbol{X}\right)^{-1} \boldsymbol{X}^\top \boldsymbol{P}_W \boldsymbol{y}.$$

Notice, however, that the OLS covariance matrix estimate from (7.28) is not the estimate we want. This estimate will be

$$\frac{\|\boldsymbol{y} - \boldsymbol{P}_W \boldsymbol{X}\tilde{\boldsymbol{\beta}}\|^2}{n-k} \left(\boldsymbol{X}^\top \boldsymbol{P}_W \boldsymbol{X}\right)^{-1}, \tag{7.29}$$

while the estimate (7.24) that was derived earlier can be written as

$$\frac{\|\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}\|^2}{n} \left(\boldsymbol{X}^\top \boldsymbol{P}_W \boldsymbol{X}\right)^{-1}. \tag{7.30}$$

These two estimates are not the same. They would be the same only if IV and OLS were identical, that is, if $\boldsymbol{X} = \boldsymbol{P}_W \boldsymbol{X}$. In addition, n would have to be replaced by $n - k$ in (7.30). The problem is that the second-stage OLS regression provides an incorrect estimate of σ^2; it uses $\boldsymbol{y} - \boldsymbol{P}_W \boldsymbol{X}\tilde{\boldsymbol{\beta}}$ rather than $\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{\beta}}$ as the vector of residuals. The second-stage residuals $\boldsymbol{y} - \boldsymbol{P}_W \boldsymbol{X}\tilde{\boldsymbol{\beta}}$ may be either too large or too small, asymptotically. Whether they are too large or too small will depend on σ^2, on the variance of the elements of $\boldsymbol{M}_W \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{P}_W \boldsymbol{X}\boldsymbol{\beta}$, and on the correlation between $\boldsymbol{M}_W \boldsymbol{X}\boldsymbol{\beta}$ and $\boldsymbol{u}$. If one actually performs 2SLS in two stages, rather than relying on a preprogrammed 2SLS or IV procedure, one must be careful to use (7.30) rather than (7.29) for the estimated covariance matrix.[2] Programs for 2SLS estimation normally replace $\boldsymbol{P}_W \boldsymbol{X}\tilde{\boldsymbol{\beta}}$ by $\boldsymbol{X}\tilde{\boldsymbol{\beta}}$ before calculating the explained sum of squares, the sum of squared residuals, the R^2, and other statistics that depend on these quantities.

There has been an enormous amount of work on the finite-sample properties of 2SLS, that is, the IV estimator $\tilde{\boldsymbol{\beta}}$. A few of the many papers in this area are Anderson (1982), Anderson and Sawa (1979), Mariano (1982), Phillips (1983), and Taylor (1983). Unfortunately, many of the results of this literature are very model-specific. One important result (Kinal, 1980) is that the m^{th} moment of the 2SLS estimator exists if and only if

$$m < l - k + 1.$$

[2] 2SLS is a special case of a regression with what Pagan (1984b, 1986) calls "generated regressors." Even when such regressions provide consistent parameter estimates, they usually provide inconsistent estimates of the covariance matrix of the parameter estimates. The inconsistency of (7.29) provides a simple example of this phenomenon.