useful when it is not clear whether it is safe to use least squares rather than instrumental variables.

Regression (7.63) deserves further comment. It has the remarkable feature that the OLS estimates of $\boldsymbol{\beta}$ are numerically identical to the IV estimates of $\boldsymbol{\beta}$ in the original model (7.01). Moreover, the estimated covariance matrices are also the same, except that the OLS estimate from (7.63) uses an inconsistent estimator for σ^2. These results are easy to obtain. Denote by $\boldsymbol{M}^*$ the orthogonal projection onto the space $\mathcal{S}^{\perp}(\boldsymbol{M}_W\boldsymbol{X}^*)$. Then, by the FWL Theorem, the OLS estimates from (7.63) must be identical to those from the regression

$$\boldsymbol{M}^*\boldsymbol{y} = \boldsymbol{M}^*\boldsymbol{X}\boldsymbol{\beta} + \text{residuals}. \tag{7.65}$$

Now

$$\boldsymbol{M}^*\boldsymbol{X} = \boldsymbol{X} - \boldsymbol{M}_W\boldsymbol{X}^*\big(\boldsymbol{X}^{*\top}\boldsymbol{M}_W\boldsymbol{X}^*\big)^{-1}\boldsymbol{X}^{*\top}\boldsymbol{M}_W\boldsymbol{X}.$$

From the fact that $\boldsymbol{M}_W\boldsymbol{X} = [\boldsymbol{M}_W\boldsymbol{X}^* \quad \mathbf{0}]$, it follows that

$$\boldsymbol{X}^{*\top}\boldsymbol{M}_W\boldsymbol{X} = \boldsymbol{X}^{*\top}\boldsymbol{M}_W\big[\boldsymbol{X}^* \quad \mathbf{0}\big].$$

Consequently, we obtain

$$\boldsymbol{M}^*\boldsymbol{X} = \boldsymbol{X} - \big[\boldsymbol{M}_W\boldsymbol{X}^* \quad \mathbf{0}\big] = \boldsymbol{X} - \boldsymbol{M}_W\boldsymbol{X} = \boldsymbol{P}_W\boldsymbol{X}.$$

Then the OLS estimate of $\boldsymbol{\beta}$ from (7.65) is seen to be

$$\big(\boldsymbol{X}^{\top}\boldsymbol{M}^*\boldsymbol{X}\big)^{-1}\boldsymbol{X}^{\top}\boldsymbol{M}^*\boldsymbol{y} = \big(\boldsymbol{X}^{\top}\boldsymbol{P}_W\boldsymbol{X}\big)^{-1}\boldsymbol{X}^{\top}\boldsymbol{P}_W\boldsymbol{y}. \tag{7.66}$$

The right-hand side of (7.66) is of course the expression for the IV or 2SLS estimate of $\boldsymbol{\beta}$, expression (7.17).

By an extension of this argument, it is easy to see that the estimated OLS covariance matrix of $\hat{\boldsymbol{\beta}}$ from (7.63) will be

$$\tilde{s}^2\big(\boldsymbol{X}^{\top}\boldsymbol{P}_W\boldsymbol{X}\big)^{-1}, \tag{7.67}$$

where $\tilde{s}^2$ denotes the OLS estimate of the error variance in (7.63). Expression (7.67) looks just like the IV covariance matrix (7.24), except that $\tilde{s}^2$ appears instead of $\tilde{\sigma}^2$. When $\boldsymbol{\eta}$ is nonzero (so that IV estimation is necessary), the variance of the errors in (7.63) will be less than σ^2. As a consequence, $\tilde{s}^2$ will be biased downward as an estimator of σ^2. Of course, it would be easy to obtain a valid estimated covariance matrix by multiplying (7.67) by $\tilde{\sigma}^2/\tilde{s}^2$.

We now return to the DWH test. A variant of this test is applicable to nonlinear models like (7.31) as well as to linear ones. The test would then be based on a variant of the Gauss-Newton regression. If the null were that the NLS estimates $\hat{\boldsymbol{\beta}}$ were consistent, an appropriate test statistic would be an asymptotic F test for $\boldsymbol{c} = \mathbf{0}$ in the GNR

$$\boldsymbol{y} - \hat{\boldsymbol{x}} = \hat{\boldsymbol{X}}\boldsymbol{b} + \boldsymbol{M}_W\hat{\boldsymbol{X}}^*\boldsymbol{c} + \text{residuals}, \tag{7.68}$$