By the same token, if the parameters of the structural model are not constant over the entire sample, then the parameters of the URF will not be constant either. Since the equations of the URF are estimated by ordinary least squares, it is very easy to test them for evidence of misspecification such as serial correlation, heteroskedasticity, and nonconstant coefficients. If they fail any of these tests, then one may reasonably conclude that the structural model is misspecified, even if one has not actually estimated it. The converse is not true, however, since these tests may well lack power, especially if only one of the structural equations is misspecified.

One additional misspecification test that should always be performed is a test of any **overidentifying restrictions**. In Section 7.8, we discussed how to test overidentifying restrictions for a single equation estimated by IV or 2SLS. Here we are interested in all of the overidentifying restrictions for the entire system. The number of degrees of freedom for the test is equal to the number of elements in the $\boldsymbol{\Pi}$ matrix of the URF, gk, minus the number of free parameters in $\boldsymbol{B}$ and $\boldsymbol{\Gamma}$ jointly. In most cases there will be some overidentifying restrictions, and in many cases there will be a large number of them. The most natural way to test these is probably to use an LR test. The restricted value of the loglikelihood function is the value of (18.30) at the FIML estimates $\hat{\boldsymbol{B}}$ and $\hat{\boldsymbol{\Gamma}}$, and the unrestricted value is

$$-\frac{ng}{2}\big(\log(2\pi)+1\big)-\frac{n}{2}\log\left|\frac{1}{n}\big(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\Pi}}\big)^{\top}\big(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\Pi}}\big)\right|, \tag{18.33}$$

where $\hat{\boldsymbol{\Pi}}$ denotes the OLS estimates of the parameters of the URF. As usual, twice the difference between the unrestricted and restricted values of the loglikelihood function will be asymptotically distributed as χ^2 with as many degrees of freedom as there are overidentifying restrictions. If one suspects that the overidentifying restrictions are violated and therefore does not want to bother estimating the structural model, one could instead use a Wald test, as suggested by Byron (1974).

We have not yet explained why the OLS estimates $\hat{\boldsymbol{\Pi}}$ are also the ML estimates. It can easily be seen from (18.33) that, in order to obtain ML estimates of $\boldsymbol{\Pi}$, we need to minimize the determinant

$$\big|(\boldsymbol{Y}-\boldsymbol{X\Pi})^{\top}(\boldsymbol{Y}-\boldsymbol{X\Pi})\big|. \tag{18.34}$$

Suppose that we evaluate this determinant at any set of estimates $\acute{\boldsymbol{\Pi}}$ not equal to $\hat{\boldsymbol{\Pi}}$. Since we can always write $\acute{\boldsymbol{\Pi}} = \hat{\boldsymbol{\Pi}} + \boldsymbol{A}$ for some matrix $\boldsymbol{A}$, (18.34) becomes

$$\begin{aligned}
&\big|(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\Pi}}-\boldsymbol{XA})^{\top}(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\Pi}}-\boldsymbol{XA})\big| \\
&= \big|(\boldsymbol{M}_X\boldsymbol{Y}-\boldsymbol{XA})^{\top}(\boldsymbol{M}_X\boldsymbol{Y}-\boldsymbol{XA})\big| \\
&= \big|\boldsymbol{Y}^{\top}\boldsymbol{M}_X\boldsymbol{Y}+\boldsymbol{A}^{\top}\boldsymbol{X}^{\top}\boldsymbol{XA}\big|.
\end{aligned} \tag{18.35}$$