The result (16.60) makes clear just what the difference is between the empirical moment evaluated at the unknown $\boldsymbol{\theta}_0$ and evaluated at the ML estimates $\hat{\boldsymbol{\theta}}$, that is, between $n^{-1/2}\boldsymbol{m}_0^{\top}\boldsymbol{\iota}$ and $n^{-1/2}\hat{\boldsymbol{m}}^{\top}\boldsymbol{\iota}$. The effect of using the estimates is an implicit orthogonal projection of the vector $\boldsymbol{m}_0$ onto the orthogonal complement of the space $\mathcal{S}(\boldsymbol{G}_0)$ associated with the model parameters. This projection is what causes the variance of the expression that we can actually calculate to be smaller than the variance of the corresponding expression based on the true parameters. The variances used in the skewness and kurtosis tests discussed in the last section can also be computed using (16.60).

We are now ready to obtain an appropriate expression for the asymptotic variance of $n^{-1/2}\hat{\boldsymbol{m}}^{\top}\boldsymbol{\iota}$. We require, as we suggested earlier, that $n^{-1/2}\boldsymbol{m}_0^{\top}\boldsymbol{\iota}$ should satisfy CLT and that, in a neighborhood of $\boldsymbol{\theta}_0$, $n^{-1}\boldsymbol{m}^{\top}(\boldsymbol{\theta})\boldsymbol{G}_i(\boldsymbol{\theta})$ should satisfy WULLN (Definition 4.17) for all $i = 1, \ldots, k$. The asymptotic variance is then clearly $\operatorname{plim}(n^{-1}\boldsymbol{m}_0^{\top}\boldsymbol{M}_G\boldsymbol{m}_0)$, which can be consistently estimated by $n^{-1}\hat{\boldsymbol{m}}^{\top}\hat{\boldsymbol{M}}_G\hat{\boldsymbol{m}}$. This suggests using the test statistic

$$\frac{n^{-1/2}\hat{\boldsymbol{m}}^{\top}\boldsymbol{\iota}}{\left(n^{-1}\hat{\boldsymbol{m}}^{\top}\hat{\boldsymbol{M}}_G\hat{\boldsymbol{m}}\right)^{1/2}} = \frac{\hat{\boldsymbol{m}}^{\top}\boldsymbol{\iota}}{\left(\hat{\boldsymbol{m}}^{\top}\hat{\boldsymbol{M}}_G\hat{\boldsymbol{m}}\right)^{1/2}}, \tag{16.61}$$

which will be asymptotically distributed as $N(0, 1)$.

The connection with the OPG regression is now evident. The test statistic (16.61) is *almost* the t statistic on the coefficient b from the following OPG regression:

$$\boldsymbol{\iota} = \hat{\boldsymbol{G}}\boldsymbol{c} + b\hat{\boldsymbol{m}} + \text{residuals}. \tag{16.62}$$

Asymptotically, the statistic (16.61) and the t statistic from (16.62) are equivalent, because the sum of squared residuals from (16.62) tends to n for large sample sizes under the null hypothesis: The regressors $\hat{\boldsymbol{G}}$ are always orthogonal to $\boldsymbol{\iota}$, and $\hat{\boldsymbol{m}}$ is orthogonal to $\boldsymbol{\iota}$ if the moment condition is satisfied. This result is very satisfactory. Without the regressor $\hat{\boldsymbol{m}}$, which is the vector that serves to define the empirical moment, regression (16.62) would be just the OPG regression associated with the original model, and the SSR would always be equal to n. Thus the OPG version of the CM test, like all the other tests we have discussed that are implemented by artificial regressions, is just a test for the significance of the coefficients on one or more test regressors.

It is now plain how to extend CM tests to a set of two or more moment conditions. One simply creates a test regressor for each of the empirical moments so as to produce an $n \times r$ matrix $\hat{\boldsymbol{R}} \equiv \boldsymbol{R}(\hat{\boldsymbol{\theta}})$, where r is the number of moment conditions. One then uses the explained sum of squares from the OPG regression

$$\boldsymbol{\iota} = \hat{\boldsymbol{G}}\boldsymbol{c} + \hat{\boldsymbol{R}}\boldsymbol{b} + \text{residuals}$$

or any other asymptotically equivalent test of the artificial hypothesis $\boldsymbol{b} = \boldsymbol{0}$. It is now clear that, as we suggested above, any test capable of being carried out by means of an OPG regression can be interpreted as a CM test.