equal to the classical Wald test statistic and share with it the property of being based exclusively on the ML estimates of the unrestricted model. Unfortunately, they also share the property of being parametrization dependent. Consider the OPG regression corresponding to the unrestricted model evaluated at $[\hat{\boldsymbol{\theta}}_1 \,\vdots\, \boldsymbol{\theta}_2^0]$, a parameter vector which, by construction, satisfies the null hypothesis. This artificial regression is

$$\boldsymbol{\iota} = \boldsymbol{G}_1(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)\boldsymbol{c}_1 + \boldsymbol{G}_2(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)\boldsymbol{c}_2 + \text{residuals}. \tag{13.93}$$

This is just a special case of the $C(\alpha)$ regression (13.91), and so any asymptotically valid test of the artificial hypothesis $\boldsymbol{c}_2 = \mathbf{0}$ based on (13.93) provides a valid Wald-like test.

LM, $C(\alpha)$, and Wald-like tests based on the OPG regression are so simple that it seems inviting to suggest that all tests other than the LR test can most conveniently be computed by means of an OPG regression. However, as is clear from (13.85) for the LM test, all tests based on the OPG regression use the outer-product-of-the-gradient estimator of the information matrix. Although this estimator has the advantage of being parametrization independent, numerous Monte Carlo experiments have shown that its finite-sample properties are almost always very different from its nominal asymptotic ones unless sample sizes are very large, often on the order of many thousand. In particular, these experiments suggest that OPG tests often have a size far in excess of their nominal asymptotic size. True null hypotheses are rejected much too often, in some especially bad cases, almost all the time. See, among others, Davidson and MacKinnon (1983a, 1985c, 1992a), Bera and McKenzie (1986), Godfrey, McAleer, and McKenzie (1988), and Chesher and Spady (1991). Although some experiments have suggested that OPG-based tests have about as much power as other variants of the classical tests *if* a way can be found to correct for their size, no one has found any easy and convenient way to perform the necessary size correction.

In view of this rather disappointing feature of the OPG regression, we must conclude this section with a firm admonition to readers to use it with great care. In most cases, it is safe to conclude that a restriction is compatible with the data if a test statistic computed using the OPG regression fails to reject the null hypothesis. But it is generally not safe to conclude that a restriction is incompatible with the data if an OPG test statistic rejects the null, at least not for samples of any ordinary size. Of course, if something is known about the properties of the particular OPG test being used, perhaps as a result of Monte Carlo experiments, one may then be able to draw conclusions from an OPG test statistic that rejects the null.

However, the OPG regression would be important even if one never actually used it to calculate test statistics. Its use in *theoretical* asymptotic calculations can make such calculations much simpler than they might otherwise be. Moreover, as we will see in the next two chapters, there exist other artificial regressions, not quite so generally applicable as the OPG one