can be used with any model estimated by maximum likelihood. The OPG regression was first used as a means of computing test statistics by Godfrey and Wickens (1981). This artificial regression, which is very easy indeed to set up for most models estimated by maximum likelihood, can be used for the same purposes as the GNR: verification of first-order conditions for the maximization of the loglikelihood function, covariance matrix estimation, one-step efficient estimation, and, of greatest immediate interest, the computation of test statistics.

Suppose that we are interested in the parametrized model (13.01). Let $\boldsymbol{G}(\boldsymbol{\theta})$ be the CG matrix associated with the loglikelihood function $\ell^n(\boldsymbol{\theta})$, with typical element

$$G_{ti}(\boldsymbol{\theta}) \equiv \frac{\partial \ell_t(\boldsymbol{\theta})}{\partial \theta_i}; \quad t = 1, \ldots, n, \;\; i = 1, \ldots, k,$$

where k is the number of elements in the parameter vector $\boldsymbol{\theta}$. Then the OPG regression associated with the model (13.01) can be written as

$$\boldsymbol{\iota} = \boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{c} + \text{residuals}. \tag{13.81}$$

Here $\boldsymbol{\iota}$ is an n–vector of which each element is unity and $\boldsymbol{c}$ is a k–vector of artificial parameters. The product of the matrix of regressors with the regressand is the gradient $\boldsymbol{g}(\boldsymbol{\theta}) \equiv \boldsymbol{G}^{\top}(\boldsymbol{\theta})\boldsymbol{\iota}$. The matrix of sums of squares and cross-products of the regressors, $\boldsymbol{G}^{\top}(\boldsymbol{\theta})\boldsymbol{G}(\boldsymbol{\theta})$, when divided by n, consistently estimates the information matrix $\mathcal{I}(\boldsymbol{\theta})$. These two features are essentially all that is required for (13.81) to be a valid artificial regression.[6] As with the GNR, the regressors of the OPG regression depend on the vector $\boldsymbol{\theta}$. Therefore, before the artificial regression is run, these regressors must be evaluated at some chosen parameter vector.

One possible choice for this parameter vector is $\hat{\boldsymbol{\theta}}$, the ML estimator for the model (13.01). In this case, the regressor matrix is $\hat{\boldsymbol{G}} \equiv \boldsymbol{G}(\hat{\boldsymbol{\theta}})$ and the artificial parameter estimates, which we will denote by $\hat{\boldsymbol{c}}$, are identically zero:

$$\hat{\boldsymbol{c}} = \big(\hat{\boldsymbol{G}}^{\top}\hat{\boldsymbol{G}}\big)^{-1}\hat{\boldsymbol{G}}^{\top}\boldsymbol{\iota} = \big(\hat{\boldsymbol{G}}^{\top}\hat{\boldsymbol{G}}\big)^{-1}\hat{\boldsymbol{g}} = \mathbf{0}.$$

Since $\hat{\boldsymbol{g}}$ here is the gradient of the loglikelihood function evaluated at $\hat{\boldsymbol{\theta}}$, the last equality above is a consequence of the first-order conditions for the maximum of the likelihood. As with the GNR, then, running the OPG regression with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ provides a simple way to test how well the first-order conditions are in fact satisfied by a set of estimates calculated by means of some computer program. The t statistics again provide the most suitable check. They should not exceed a number around 10^{-2} or 10^{-3} in absolute value if a good approximation to the maximum has been found.

[6] Precise conditions for a regression to be called "artificial" are provided by Davidson and MacKinnon (1990); see Section 14.4.