so and is probably to be preferred, since the factor of s^2 in (15.23) simply introduces additional randomness into the estimate of the covariance matrix.

As usual, the covariance matrix of $\hat{\boldsymbol{\beta}}$ can also be estimated as minus the inverse of the numerical Hessian or as the inverse of the outer product of the CG matrix, $\hat{\boldsymbol{G}}^\top\hat{\boldsymbol{G}}$. In the case of the logit model, minus the numerical Hessian is actually equal to the estimated information matrix $\boldsymbol{X}^\top\hat{\boldsymbol{\Psi}}\boldsymbol{X}$, because

$$\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial\beta_i\partial\beta_j} = \frac{\partial}{\partial\beta_j}\left(\sum_{t=1}^{n}\big(y_t - \Lambda(\boldsymbol{X}_t\boldsymbol{\beta})\big)X_{ti}\right) = -\sum_{t=1}^{n}\lambda(\boldsymbol{X}_t\boldsymbol{\beta})X_{ti}X_{tj}.$$

However, in the case of most other binary response models, including the probit model, minus the Hessian will differ from, and generally be more complicated than, the information matrix.

Like all artificial regressions, the BRMR is particularly useful for hypothesis testing. Suppose that $\boldsymbol{\beta}$ is partitioned as $[\boldsymbol{\beta}_1 \,\vdots\, \boldsymbol{\beta}_2]$, where $\boldsymbol{\beta}_1$ is a $(k-r)$–vector and $\boldsymbol{\beta}_2$ is an r–vector. If $\tilde{\boldsymbol{\beta}}$ denotes the vector of ML estimates subject to the restriction that $\boldsymbol{\beta}_2 = \boldsymbol{0}$, we can test that restriction by running the BRMR

$$\tilde{V}_t^{-1/2}(y_t - \tilde{F}_t) = \tilde{V}_t^{-1/2}\tilde{f}_t\boldsymbol{X}_{t1}\boldsymbol{b}_1 + \tilde{V}_t^{-1/2}\tilde{f}_t\boldsymbol{X}_{t2}\boldsymbol{b}_2 + \text{residual}, \tag{15.24}$$

where $\tilde{F}_t \equiv F(\boldsymbol{X}_t\tilde{\boldsymbol{\beta}})$, $\tilde{f}_t \equiv f(\boldsymbol{X}_t\tilde{\boldsymbol{\beta}})$, and $\tilde{V}_t \equiv V(\boldsymbol{X}_t\tilde{\boldsymbol{\beta}})$. Here $\boldsymbol{X}_t$ has been partitioned into two vectors, $\boldsymbol{X}_{t1}$ and $\boldsymbol{X}_{t2}$, corresponding to the partitioning of $\boldsymbol{\beta}$. The regressors that correspond to $\boldsymbol{\beta}_1$ are orthogonal to the regressand, while those that correspond to $\boldsymbol{\beta}_2$ are not. All the usual test statistics for $\boldsymbol{b}_2 = \boldsymbol{0}$ are valid. However, in contrast to the case of the Gauss-Newton regression, there is no particular reason to use an F test, because there is no variance parameter to estimate. The best test statistic to use in finite samples, according to Monte Carlo results obtained by Davidson and MacKinnon (1984b), is probably the explained sum of squares from regression (15.24). It will be asymptotically distributed as $\chi^2(r)$ under the null hypothesis. Note that nR^2 will not be equal to the explained sum of squares in this case, because the total sum of squares will not be equal to n.

In one very special case, the BRMR (15.24) becomes extremely simple. Suppose the null hypothesis is that all the slope coefficients are zero. In this case, $\boldsymbol{X}_{t1}$ is just unity, $\boldsymbol{X}_t\tilde{\boldsymbol{\beta}} = \tilde{\beta}_1 = F^{-1}(\bar{y})$, and, in obvious notation, regression (15.24) becomes

$$\bar{V}^{-1/2}(y_t - \bar{F}) = \bar{V}^{-1/2}\bar{f}b_1 + \bar{V}^{-1/2}\bar{f}\boldsymbol{X}_{t2}\boldsymbol{b}_2 + \text{residual}.$$

Neither subtracting a constant from the regressand nor multiplying the regressand and regressors by a constant has any effect on the F statistic for $\boldsymbol{b}_2 = \boldsymbol{0}$. Thus it is clear that we can test the all-slopes-zero hypothesis simply by calculating an F statistic for $\boldsymbol{c}_2 = \boldsymbol{0}$ in the linear regression

$$\boldsymbol{y} = c_1 + \boldsymbol{X}_2\boldsymbol{c}_2 + \text{residuals}.$$