Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE @ DIRECT®

Journal of Statistical Planning and Inference 124 (2004) 423–437

journal of statistical planning and inference

www.elsevier.com/locate/jspi

An improved test for heteroskedasticity using adjusted modified profile likelihood inference

Silvia L.P. Ferrari[a,*], Audrey H.M.A. Cysneiros[b], Francisco Cribari-Neto[b]

[a] *Departamento de Estatística, IME, Universidade de São Paulo, Caixa Postal 66281, São Paulo/SP 05311-970, Brazil*

[b] *Departamento de Estatística, CCEN, Universidade Federal de Pernambuco, Cidade Universitária, Recife/PE 50740-540, Brazil*

Received 12 October 2002; accepted 7 April 2003

Abstract

This paper addresses the issue of testing for heteroskedasticity in linear regression models. We derive a Bartlett adjustment to the modified profile likelihood ratio test (J. Roy. Statist. Soc. B 49 (1987) 1) for heteroskedasticity in the normal linear regression model. Our results generalize those in Ferrari and Cribari-Neto (Statist. Probab. Lett. 57 (2002) 353), since they allow for a vector-valued structure for the parameter that defines the skedastic function. Monte Carlo evidence shows that the proposed test displays reliable finite-sample behavior, outperforming the original likelihood ratio test, the Bartlett-corrected likelihood ratio test, and the modified profile likelihood ratio test.

Keywords: Bartlett correction; Heteroskedasticity; Likelihood ratio test; Nuisance parameters; Profile likelihood

1. Introduction

Linear regression models are oftentimes used to model the behavior of a variable of interest conditional on a set of explanatory variables. In practical applications, it is common for these models to include a heteroskedastic structure, thus indicating that the conditional variances are not constant across observations. Since the modeling strategies are different when such variances are not constant, it is important to first test

* Corresponding author.
E-mail addresses: sferrari@ime.usp.br (S.L.P. Ferrari), audrey@de.ufpe.br (A.H.M.A. Cysneiros), cribari@ufpe.br (F. Cribari-Neto).

0378-3758/$ - see front matter
doi:10.1016/S0378-3758(03)00210-6

whether heteroskedasticity is present in the data. The most commonly used tests for heteroskedasticity are based on first-order asymptotics, since they rely on a large sample approximation. The asymptotic approximation used to perform the tests, however, may not deliver accurate inference when the sample size is not large. In particular, likelihood ratio tests for constancy of variances in linear regression models tend to be severely oversized in samples of typical size. It is thus important to develop inference strategies that are more reliable and display superior finite-sample behavior. Cordeiro (1993) proposed a Bartlett adjustment to the likelihood ratio test for heteroskedasticity, where the test statistic is modified using an adjustment factor that typically reduces finite-sample size discrepancies. The original likelihood ratio statistic is distributed as χ^2 up to an error of order n^{-1} whereas the transformed statistic follows the same limiting distribution up to an error of order n^{-2}. Simonoff and Tsai (1994) used an approximation to the modified profile likelihood inference approach proposed by Cox and Reid (1987) to develop a test for heteroskedasticity aiming at reducing the effect of nuisance parameters on the resulting inference. The proposed test typically delivers more reliable inference in finite samples, but the test statistic is still χ^2 distributed, under the null hypothesis, up to an error of order n^{-1}. Ferrari and Cribari-Neto (2002) have Bartlett-corrected this test to obtain an adjusted test with superior finite-sample behavior. Their results, however, have an important limitation: they are only valid for situations where the parameter that defines the heteroskedastic behavior of the data is scalar. Since in most applications practitioners use variance specifications that are based on more than just one covariate, it is important to develop second-order accurate adjusted modified profile likelihood inference in wider generality. This is the goal of the present paper.

The paper unfolds as follows. Section 2 presents the model of interest and the aforementioned tests for heteroskedasticity. Section 3 develops the main result in the article, namely the second-order accurate inference based on adjusted modified profile likelihood methods. Numerical evidence on the finite-sample behavior of the different tests is presented in Section 4. The numerical evidence favors the test proposed in this paper. Section 5 considers applications to real data. Concluding remarks are given in Section 6.

2. Some tests for heteroskedasticity

The model of interest is

$$y = X\beta + u, \tag{1}$$

where y is an n-vector of observations on the dependent variable, X is an $n \times k$ nonrandom matrix of covariates with full column rank, β is a k-vector of unknown regression parameters, and u is an n-vector of random errors. It is assumed that $u \sim \mathcal{N}(0, \sigma^2 W)$, where W is an $n \times n$ diagonal matrix with the lth entry given by $w_l = w(z_l, \delta) > 0$, z_l' being the lth row of the $n \times p$ matrix Z of variance covariates, σ^2 is a strictly positive and finite unknown constant, and δ is a $p \times 1$ unknown parameter vector. The desired inference is obtained by testing the null hypothesis $\mathscr{H}_0$: $\delta = \delta_0$

(homoskedasticity), where δ_0 is a given p-vector of constants such that $w(z_l, \delta_0) = 1$ for $l = 1, \ldots, n$, against a two-sided alternative hypothesis. The number of parameters of interest is therefore p, and the number of nuisance parameters is $k+1$. The resulting log-likelihood function is

$$L = L(y; \delta, \beta, \sigma^2) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2}\sum_{l=1}^{n}\log w(z_l, \delta)$$

$$-\frac{1}{2}\,\sigma^2(y - X\beta)' W^{-1}(y - X\beta) + \text{constant}.$$

The maximum likelihood estimator (MLE) of δ can be obtained by maximizing the profile log-likelihood function

$$L_p(y; \delta) = L(y; \delta, \hat{\beta}_\delta, \hat{\sigma}^2_\delta), \tag{2}$$

where

$$\hat{\beta}_\delta = (X'W^{-1}X)^{-1}X'W^{-1}y \quad \text{and} \quad \hat{\sigma}^2_\delta = \frac{1}{n}(y - X\hat{\beta}_\delta)'W^{-1}(y - X\hat{\beta}_\delta).$$

That is,

$$L_p(y; \delta) = -\frac{n}{2}\log(\hat{\sigma}^2_\delta) - \frac{1}{2}\sum_{l=1}^{n}\log w(z_l, \delta) + \text{constant}.$$

The likelihood ratio statistic for the test at hand is

$$\text{LR} = -2\{L_p(y; \delta_0) - L_p(y; \hat{\delta})\}, \tag{3}$$

where $\hat{\delta}$ is the MLE of δ. The null distribution of LR in large samples may be approximated by a χ^2_p distribution. However, it is well known that this approximation can be poor if the sample size is not large enough. This shortcoming can be circumvented by modifying the test statistic using a Bartlett correction in order to improve on the approximation. The Bartlett-corrected likelihood ratio statistic

$$\text{LR}^* = \frac{\text{LR}}{1 + c/p},$$

where c is a suitably chosen constant of order n^{-1} such that $E(\text{LR}) = p\{(1 + c/p) + \text{O}(n^{-3/2})\}$, has a χ^2_p null distribution an error of order n^{-2} (see Barndorff-Nielsen and Hall, 1988). The gain in accuracy stems from the fact that $\text{pr}(\text{LR}^* > x_\alpha) = \alpha + \mathcal{O}(n^{-2})$ whereas for the unmodified statistic $\text{pr}(\text{LR} > x_\alpha) = \alpha + \mathcal{O}(n^{-1})$, where x_α is the $1-\alpha$ upper point from a χ^2_p distribution. A general expression for c can be found in Lawley (1956); see Cribari-Neto and Cordeiro (1996). The Bartlett adjustment factor needs to be tailored for each application of interest, and Cordeiro (1993) has obtained a closed form expression for such a constant for the test for heteroskedasticity in linear regression. His main result is reproduced by Ferrari and Cribari-Neto (2002, p. 355).

As indicated earlier, Simonoff and Tsai (1994) have used the modified profile likelihood approach proposed by Cox and Reid (1987) to obtain an alternative homoskedasticity test statistic. To that end, it is required that δ be orthogonal to the remaining parameters. For the model of interest here, the transformation $(\delta', \beta', \sigma^2)' \to (\delta', \beta', \gamma)'$, with

$$\sigma^2 = \frac{\gamma}{(\prod_{l=1}^{n} w_l)^{1/n}},$$

delivers the desired orthogonality. The log-likelihood function for the reparameterized model is given by

$$L^* = L^*(y; \delta, \beta, \gamma) = -\frac{n}{2}\log\gamma - \frac{1}{2}\gamma(y - X\beta)'V(y - X\beta) + \text{constant},$$

where $V = \text{diag}\{v_1, \ldots, v_n\}$, with

$$v_l = v_l(\delta) = w_l^{-1}\left(\prod_{m=1}^{n} w_m\right)^{1/n}. \tag{4}$$

The corresponding profile log-likelihood function for δ is

$$L_p^* = L_p^*(y; \delta) = -\frac{n}{2}\log\hat{\gamma}_\delta + \text{constant},$$

where $\hat{\gamma}_\delta = \hat{\sigma}_\delta^2(\prod_{l=1}^{n} w_l)^{1/n}$.

The modified profile log-likelihood function for δ is (Cox and Reid, 1987)

$$L_{mp}^*(y; \delta) = L_p^*(y; \delta) - \frac{1}{2}\log[\det(j^*(\delta; \hat{\beta}_\delta, \hat{\gamma}_\delta))],$$

where $j^*(\delta; \hat{\beta}_\delta, \hat{\gamma}_\delta)$ is the block of the observed information matrix corresponding to the nuisance parameters $(\beta', \gamma)'$, in the reparameterized model, evaluated at $(\delta; \hat{\beta}_\delta, \hat{\gamma}_\delta)$. It is easy to show that

$$L_{mp}^*(y; \delta) = L_p^*(y; \delta) - \frac{1}{2}\log\left\{\frac{n}{2\hat{\gamma}_\delta^{k+2}}\det(X_m'X_m)\right\},$$

where $X_m = G^{-1/2}X$, $G = G(\delta)$ being an $n \times n$ diagonal matrix whose lth diagonal element is given by $w(z_l, \delta)/\{\prod_{m=1}^{n} w(z_m, \delta)\}^{1/n}$.

The likelihood ratio statistic based on the modified profile likelihood function for the test of $\mathscr{H}_0$ vs. $\mathscr{H}_1$ is given by

$$\text{LR}_m = -2\{L_{mp}^*(y; \delta_0) - L_{mp}^*(y; \tilde{\delta})\},$$

where $\tilde{\delta}$ is the value of δ that maximizes $L_{mp}^*(y; \delta)$. Note that LR_m requires maximization of the modified profile likelihood function. It is easy to show that the resulting test statistic can be written as

$$\text{LR}_m = \frac{n-k-2}{n}\text{LR} + \log\left\{\frac{\det(X'X)}{\det(\tilde{X}_m'\tilde{X}_m)}\right\},$$

where $\tilde{X}_m = \tilde{G}^{-1/2}X$. Here, $\tilde{G} = G(\tilde{\delta})$. Simonoff and Tsai (1994) used an alternative form of the test statistic that avoids such maximization by replacing $\tilde{\delta}$ by $\hat{\delta}$ in the expression for LR_m. Their test statistic is the same as the one given above, but with $\tilde{X}_m$ replaced by $\hat{X}_m$, where $\hat{X}_m = \hat{G}^{-1/2}X$, with $\hat{G} = G(\hat{\delta})$.

3. Second-order accurate profile likelihood inference

The modified profile likelihood ratio test presented in Section 2 typically delivers improved inference relative to the original likelihood ratio test, but both tests are based on first-order asymptotics. It is possible to develop a Bartlett adjustment to such a test in order to obtain a new test that is second-order accurate. This will be done in this section using the results in DiCiccio and Stern (1994). We find a correction factor c_m that defines the transformed statistic

$$\mathrm{LR}_m^* = \frac{\mathrm{LR}_m}{1 + c_m/p}$$

such that $\mathrm{pr}(\mathrm{LR}_m^* > x_\alpha) = \alpha + \mathcal{O}(n^{-2})$.

The notation used can be summarized as follows, where indices $a, b, \ldots$ range over $1, \ldots, p$, and indices $r, s, t, \ldots$ range over $1, \ldots, p+k+1$. Let $\theta = (\delta', \beta', \gamma)'$, θ^r denoting the rth element of θ, $\lambda_{rs} = E(\partial^2 L/\partial\theta^r\partial\theta^s)$, $\lambda_{rst} = E(\partial^3 L/\partial\theta^r\partial\theta^s\partial\theta^t)$, etc. Define $(\lambda_{rs})_t = \partial\lambda_{rs}/\partial\theta^t$, $(\lambda_{rs})_{tu} = \partial^2\lambda_{rs}/\partial\theta^t\partial\theta^u$, etc. The (r,s) element of Fisher's information matrix is thus given by $-\lambda_{rs}$, and the corresponding element of its inverse is $-\lambda^{rs}$. Additionally, $\tau^{rs} = \lambda^{ra}\lambda^{sb}\sigma_{ab}$, where (σ_{ab}) is the $p \times p$ matrix inverse of (λ^{ab}), and $v^{rs} = \lambda^{rs} - \tau^{rs}$. Note that the entries of the $(p+k+1) \times (p+k+1)$ matrix (v^{rs}) are all zero except for its lower right-hand $(k+1) \times (k+1)$ submatrix which is the inverse of Fisher's information matrix for the nuisance parameters (β', γ) keeping δ fixed. The implicit summation convention is used throughout, i.e., indices repeated as subscripts and superscripts indicate summation over the appropriate range.

DiCiccio and Stern (1994, p. 404, Eq. (25)) obtained a general expression for c_m; it can be written as

$$\begin{aligned}
c_m = {} & \tfrac{1}{4}\tau^{ru}\tau^{st}\lambda_{rstu} - \lambda^{ru}\tau^{st}(\lambda_{rst})_u + (\lambda^{ru}\lambda^{st} - v^{ru}v^{st})(\lambda_{rs})_{tu} \\
& - (\tfrac{1}{4}\lambda^{ru}\tau^{st}\tau^{vw} + \tfrac{1}{2}\lambda^{ru}\tau^{sw}\tau^{tv} - \tfrac{1}{3}\tau^{ru}\tau^{sw}\tau^{tv})\lambda_{rst}\lambda_{uvw} \\
& + (\lambda^{ru}\tau^{st}\lambda^{vw} + \lambda^{ru}\lambda^{sw}\lambda^{tv} - v^{ru}\lambda^{sw}v^{tv})\lambda_{rst}(\lambda_{uv})_w \\
& - (\lambda^{ru}\lambda^{st}\lambda^{vw} - v^{ru}v^{st}v^{vw} + \lambda^{ru}\lambda^{sw}\lambda^{tv} - v^{ru}v^{sw}v^{tv})(\lambda_{rs})_t(\lambda_{uv})_w.
\end{aligned}$$

In our setup, i.e., where the parameter vector is written in partitioned form as $(\delta', \beta', \gamma)'$ and γ is scalar, we have that $v^{ab} = v^{a\gamma} = v^{ai} = 0$, $\tau^{ab} = \lambda^{ab}$, $\tau^{a\gamma} = \lambda^{a\gamma}$, $\tau^{ai} = \lambda^{ai}$, and $\tau^{i\gamma} = \lambda^{i\gamma}$, where γ is used to represent the index corresponding to the last component of the parameter vector $(\delta', \beta', \gamma)'$, and $i, j, \ldots$ range over $p+1, \ldots, p+k$. Using the orthogonality between δ, β and γ, it can be shown that $\tau^{ij} = \tau^{\gamma\gamma} = \lambda_{ai} = \lambda_{a\gamma} = \lambda_{i\gamma} = \lambda^{ai} = \lambda^{a\gamma} = \lambda^{i\gamma} = 0$, $v^{ij} = \lambda^{ij}$ and $\lambda^{\gamma\gamma} = v^{\gamma\gamma}$. After long algebra, we obtain the following expression for c_m:

$$\begin{aligned}
c_m = {} & \tfrac{1}{4}\lambda^{ab}\lambda^{cd}\lambda_{abcd} - \lambda^{ab}\lambda^{cd}(\lambda_{acd})_b + \lambda^{ab}\lambda^{cd}(\lambda_{ac})_{db} - \lambda^{ij}\lambda^{ab}(\lambda_{iab})_j \\
& - \lambda^{\gamma\gamma}\lambda^{ab}(\lambda_{ab\gamma})_\gamma - (\tfrac{1}{4}\lambda^{ab}\lambda^{cd}\lambda^{ef} + \tfrac{1}{2}\lambda^{ab}\lambda^{cf}\lambda^{de} - \tfrac{1}{3}\lambda^{ab}\lambda^{cf}\lambda^{de})\lambda_{acd}\lambda_{bef} \\
& + (\lambda^{ab}\lambda^{cd}\lambda^{ef} + \lambda^{ab}\lambda^{cf}\lambda^{de})\lambda_{acd}(\lambda_{be})_f - (\lambda^{ab}\lambda^{cd}\lambda^{ef} + \lambda^{ab}\lambda^{cf}\lambda^{de})(\lambda_{ac})_d(\lambda_{be})_f
\end{aligned}$$

$$- (\tfrac{1}{4}\lambda^{ij}\lambda^{ab}\lambda^{cd} + \tfrac{1}{2}\lambda^{ij}\lambda^{ad}\lambda^{bc})\lambda_{iab}\lambda_{jcd} + \lambda^{ij}\lambda^{ab}\lambda^{kl}\lambda_{iab}(\lambda_{jk})_l$$
$$- (\tfrac{1}{4}\lambda^{\gamma\gamma}\lambda^{ab}\lambda^{cd} + \tfrac{1}{2}\lambda^{\gamma\gamma}\lambda^{ad}\lambda^{bc})\lambda_{ab\gamma}\lambda_{cd\gamma} + (\lambda^{\gamma\gamma}\lambda^{ab}\lambda^{\gamma\gamma})\lambda_{ab\gamma}(\lambda_{\gamma\gamma})_\gamma. \quad (5)$$

It is noteworthy that the expression we derived, expression (5), can be used to obtain closed-form expressions for the Bartlett adjustment factor in any class of models that uses the partition of the parameter vector as here and where orthogonality holds. In that sense, Eq. (5) is quite general and can be used to Bartlett-adjust modified profile likelihood ratio tests in classes of models other than the one we focus in this paper.

In what follows, we shall consider the framework described in the previous section. It can be easily shown that

$$\frac{\partial L^*}{\partial \delta^a} = -\frac{1}{2\gamma}(y - X\beta)' V_{R_1}(y - X\beta).$$

The expression for V_{R_1} follows from

$$V_{R_m} = \mathrm{diag}\left\{ \frac{\partial^m v_1}{\partial\delta^{a_1}\cdots\partial\delta^{a_m}}, \ldots, \frac{\partial^m v_n}{\partial\delta^{a_1}\cdots\partial\delta^{a_m}} \right\}$$

with $m = 1$, where $v_1, \ldots, v_n$ are given in (4). Using the fact that if $E(z) = \mu$ and $\mathrm{cov}(z) = \Sigma$ for an $n \times 1$ random vector z, then $E(z'Az) = \mathrm{tr}(A\Sigma) + \mu' A\mu$, A being an $n \times n$ nonrandom matrix, we obtain

$$E\left(\frac{\partial L^*}{\partial \delta^a}\right) = -\frac{1}{2}\,\mathrm{tr}(V_{R_1}V^{-1}) = -\frac{1}{2}\sum_{l=1}^{n}\frac{v_{l_a}}{v_l},$$

where $v_{l_a} = \partial v_l/\partial\delta^a$. It follows from regularity conditions that the above expected value equals zero, so the w's must satisfy $\sum_{r=1}^{n} v_{l_a}/v_l = 0$. Note that

$$E\left(\frac{\partial^2 L^*}{\partial\delta^a\partial\gamma}\right) = \frac{1}{2\gamma^2}\,E\{(y - X\beta)'V_{R_1}(y - X\beta)\} = \frac{1}{2\gamma}\,\mathrm{tr}(V_{R_1}V^{-1}) = \frac{1}{2\gamma}\sum_{l=1}^{n}\frac{v_{l_a}}{v_l} = 0,$$

so that orthogonality between δ and γ is preserved, as desired. The second derivative of L^* with respect to the elements of δ is

$$\frac{\partial^2 L^*}{\partial\delta^a\partial\delta^b} = -\frac{1}{2\gamma}\,(y - X\beta)'V_{R_2}(y - X\beta).$$

Therefore,

$$\lambda_{ab} = -\frac{1}{2}\,\mathrm{tr}(V_{R_2}V^{-1}) = -\frac{1}{2}\sum_{l=1}^{n}\frac{v_{l_{ab}}}{v_l},$$

where $v_{l_{ab}} = \partial^2 v_l/(\partial\delta^a\partial\delta^b)$; in what follows, $v_{l_{abc}} = \partial^3 v_l/(\partial\delta^a\partial\delta^b\partial\delta^c)$ and $v_{l_{abcd}} = \partial^4 v_l/(\partial\delta^a\partial\delta^b\partial\delta^c\partial\delta^d)$. After some algebra, we also obtain

$$\lambda_{abc} = -\frac{1}{2}\sum_{l=1}^{n}\frac{v_{l_{abc}}}{v_l}, \qquad \lambda_{abcd} = -\frac{1}{2}\sum_{l=1}^{n}\frac{v_{l_{abcd}}}{v_l}, \qquad \lambda_{ab\gamma} = \frac{1}{2\gamma}\sum_{l=1}^{n}\frac{v_{l_{ab}}}{v_l},$$

$$(\lambda_{ab\gamma})_\gamma = -\frac{1}{2\gamma^2}\sum_{l=1}^{n}\frac{v_{l_{ab}}}{v_l}, \qquad \lambda_{\gamma\gamma} = -\frac{n}{2\gamma^2}, \qquad \lambda^{\gamma\gamma} = -\frac{2\gamma^2}{n}, \qquad (\lambda_{\gamma\gamma})_\gamma = \frac{n}{\gamma^3},$$

$$(\lambda_{abc})_d = -\frac{1}{2}\sum_{l=1}^{n}\frac{\partial}{\partial\delta^d}\left(\frac{v_{l_{abc}}}{v_l}\right), \qquad (\lambda_{ab})_{cd} = -\frac{1}{2}\sum_{l=1}^{n}\frac{\partial^2}{\partial\delta^c\partial\delta^d}\left(\frac{v_{l_{ab}}}{v_l}\right),$$

$$(\lambda_{ab})_c = -\frac{1}{2}\sum_{l=1}^{n}\frac{\partial}{\partial\delta^c}\left(\frac{v_{l_{ab}}}{v_l}\right), \qquad \lambda_{iab} = 0.$$

It is now possible to further simplify the expression given for c_m:

$$\begin{aligned} c_m = {} & \lambda^{ab}\lambda^{cd}\{\tfrac{1}{4}\lambda_{abcd} - (\lambda_{acd})_b + (\lambda_{ac})_{db} - \tfrac{1}{4}\lambda^{\gamma\gamma}\lambda_{ab\gamma}\lambda_{cd\gamma}\} \\ & -\lambda^{ab}\lambda^{cd}\lambda^{ef}\{\tfrac{1}{4}\lambda_{acd}\lambda_{bef} + \lambda_{acd}(\lambda_{be})_f - (\lambda_{ac})_d(\lambda_{be})_f\} \\ & -\lambda^{ab}\lambda^{cf}\lambda^{de}\{\tfrac{1}{6}\lambda_{acd}\lambda_{bef} + \lambda_{acd}(\lambda_{be})_f - (\lambda_{ac})_d(\lambda_{be})_f\} \\ & -\lambda^{ab}\{\lambda^{\gamma\gamma}(\lambda_{ab\gamma})_\gamma - (\lambda^{\gamma\gamma})^2\lambda_{ab\gamma}(\lambda_{\gamma\gamma})_\gamma\} \\ & -\tfrac{1}{2}\lambda^{ad}\lambda^{bc}\lambda^{\gamma\gamma}\lambda_{ab\gamma}\lambda_{cd\gamma}. \end{aligned} \tag{6}$$

In what follows, we shall consider the case of multiplicative heteroskedasticity, i.e., the special case where $w_l = \exp\{z_l'\delta\}$. Here,

$$v_l = \exp\{-(z_l-\bar{z})'\delta\}, \qquad v_{l_a} = -(z_l-\bar{z})_a\exp\{-(z_l-\bar{z})'\delta\} = -(z_l-\bar{z})_a v_l,$$

$$v_{l_{ab}} = (z_l-\bar{z})_{ab}v_l, \quad v_{l_{abc}} = -(z_l-\bar{z})_{abc}v_l \quad \text{and} \quad v_{l_{abcd}} = (z_l-\bar{z})_{abcd}v_l,$$

with $\bar{z} = (\bar{z}_1,\ldots,\bar{z}_p)'$ and $\bar{z}_a = n^{-1}\sum_{l=1}^{n} z_{la}$ for $a = 1,\ldots,p$. Also, $(z_l-\bar{z})_a = z_{la}-\bar{z}_a$, $(z_l-\bar{z})_{ab} = (z_{la}-\bar{z}_a)(z_{lb}-\bar{z}_b)$, and so on. We have that

$$\lambda_{abcd} = -\frac{1}{2}\sum_{l=1}^{n}(z_l-\bar{z})_{abcd}, \quad \lambda_{abc} = \frac{1}{2}\sum_{l=1}^{n}(z_l-\bar{z})_{abc}, \quad \lambda_{ab} = -\frac{1}{2}\sum_{l=1}^{n}(z_l-\bar{z})_{ab},$$

$$\lambda_{ab\gamma} = \frac{1}{2\gamma}\sum_{l=1}^{n}(z_l-\bar{z})_{ab}, \qquad (\lambda_{ab\gamma})_\gamma = -\frac{1}{2\gamma^2}\sum_{l=1}^{n}(z_l-\bar{z})_{ab},$$

$$\lambda^{\gamma\gamma} = -\frac{2\gamma^2}{n}, \qquad (\lambda_{acd})_b = (\lambda_{ac})_{db} = (\lambda_{be})_f = 0.$$

Plugging these quantities into Eq. (6) it is possible to write c_m as

$$\begin{aligned} c_m = {} & -\frac{1}{8}\sum_{l=1}^{n}(z_l-\bar{z})_a\lambda^{ab}(z_l-\bar{z})_b(z_l-\bar{z})_c\lambda^{cd}(z_l-\bar{z})_d \\ & +\frac{1}{8n}\sum_{l=1}^{n}(z_l-\bar{z})_a\lambda^{ab}(z_l-\bar{z})_b\sum_{m=1}^{n}(z_m-\bar{z})_c\lambda^{cd}(z_m-\bar{z})_d \end{aligned}$$

$$- \frac{1}{16} \sum_{l=1}^{n} \sum_{m=1}^{n} (z_l - \bar{z})_a \lambda^{ab} (z_l - \bar{z})_b (z_l - \bar{z})_c \lambda^{cd} (z_m - \bar{z})_d (z_m - \bar{z})_e \lambda^{ef} (z_m - \bar{z})_f$$

$$- \frac{1}{24} \sum_{l=1}^{n} \sum_{m=1}^{n} (z_l - \bar{z})_a \lambda^{ab} (z_m - \bar{z})_b (z_l - \bar{z})_c \lambda^{cd} (z_m - \bar{z})_d (z_l - \bar{z})_e \lambda^{ef} (z_m - \bar{z})_f$$

$$+ \frac{1}{n} \sum_{l=1}^{n} (z_l - \bar{z})_a \lambda^{ab} (z_l - \bar{z})_b$$

$$+ \frac{1}{4n} \sum_{l=1}^{n} \sum_{m=1}^{n} (z_l - \bar{z})_a \lambda^{ab} (z_m - \bar{z})_b (z_l - \bar{z})_c \lambda^{cd} (z_m - \bar{z})_d .$$

Let $H = \{h_{lm}\} = (Z - \bar{Z})[(Z - \bar{Z})'(Z - \bar{Z})]^{-1}(Z - \bar{Z})'$, with $(Z - \bar{Z}) = (z_1 - \bar{z}, \ldots, z_n - \bar{z})'$. Note that $(z_l - \bar{z})_a \lambda^{ab} (z_m - \bar{z})_b = -2h_{lm}$. It is then possible to reduce the above expression for c_m to

$$c_m = -\frac{1}{2} \sum_{l=1}^{n} h_{ll}^2 + \frac{1}{2n} \left(\sum_{l=1}^{n} h_{ll} \right)^2 + \frac{1}{2} \sum_{l=1}^{n} \sum_{m=1}^{n} h_{ll} h_{lm} h_{mm}$$

$$+ \frac{1}{3} \sum_{l=1}^{n} \sum_{m=1}^{n} h_{lm}^3 - \frac{2}{n} \sum_{l=1}^{n} h_{ll} + \frac{1}{n} \sum_{l=1}^{n} \sum_{m=1}^{n} h_{lm}^2 .$$

It is noteworthy that $\sum_{l=1}^{n} h_{ll} = \mathrm{tr}(H) = p$. Therefore, c_m can be written in matrix form as

$$c_m = -\frac{1}{2} \mathrm{tr}(H_d H_d) + \frac{1}{2n} p^2 + \frac{1}{2} \imath' H_d H H_d \imath + \frac{1}{3} \imath' H^{(3)} \imath - \frac{2}{n} p + \frac{1}{n} \imath' H^{(2)} \imath, \tag{7}$$

where $\imath$ is an n-vector of ones, $H_d = \mathrm{diag}\{h_{11}, \cdots, h_{nn}\}$, $H^{(2)} = (h_{lm}^2)$ and $H^{(2)} = (h_{lm}^3)$.

Eq. (7) gives a simple matrix expression for the Bartlett adjustment factor. The expression only involves simple matrix operations, and can be easily implemented into any computer algebra system or statistical software that can carry out simple linear algebra operations. It can thus be used with minimal effort by practitioners. We also note that c_m in (7) depends only on the matrix Z of covariates used to model the fluctuations in the conditional variance of the response, on the number of unknown parameters in the skedastic function and on the number of observations. In particular, it is important to note that the Bartlett adjustment factor does not depend on the number of nuisance parameters (nor does it depend on any unknown parameters). This can be interpreted in light of the fact that the Cox and Reid adjustment to the profile likelihood aims at reducing the impact of nuisance parameters on the resulting inference, and this adjustment was used prior to the Bartlett correction of the test statistic. We note that expression (7) generalizes the result in Ferrari and Cribari-Neto (2002, p. 358), which only holds for $p = 1$, as indicated earlier. It is also noteworthy that they have applied the Bartlett correction to the test statistic used by Simonoff and Tsai (1994), which, as indicated earlier, is an approximation to the exact form of the test statistic obtained from the modified profile log-likelihood function of Cox and Reid (1987). Here, we shall apply the Bartlett adjustment to the exact form of the modified profile likelihood ratio test.

4. Numerical evidence

The simulation results are based on the linear regression model

$$y_l = \beta_1 + \beta_2 x_{l2} + \cdots + \beta_k x_{lk} + u_l, \qquad l = 1, \ldots, n,$$

where $u_l \sim \mathcal{N}(0, \sigma^2 \exp\{\delta_1 z_{l1} + \cdots + \delta_p z_{lp}\})$, and $\text{cov}(u_l, u_m) = 0$ for all $l \neq m$. When $\delta_1 = \cdots = \delta_p = 0$ the model is homoskedastic, with heteroskedasticity arising when $\min\{|\delta_1|, \ldots, |\delta_p|\} \neq 0$. The simulations were carried out for different values of p and k. The covariates $x_2, \ldots, x_k$ were generated as independent draws from a standard uniform distribution $\mathcal{U}(0, 1)$. When $p < k$, the matrix Z was formed using columns $2, \ldots, p+1$ of X; when $p \geqslant k$, the extra columns of Z were created using independent draws from the $\mathcal{U}(0, 1)$ distribution.

We shall report the null rejection rates of the original likelihood ratio test (LR), its Bartlett-corrected version (LR*), the modified profile likelihood ratio test (LR$_m$), and its Bartlett-corrected version (LR$_m^*$) for the test of the null hypothesis $\mathscr{H}_0$: $\delta_1 = \cdots = \delta_p = 0$. The number of replications was set at 10,000 and the following levels of significance were considered: $\alpha = 0.100, 0.050, 0.010, 0.005$. The simulations were performed using the `Ox` matrix programming language (Doornik, 2001). The nonlinear maximizations of the relevant log-likelihoods were carried out using the BFGS quasi-Newton algorithm, which is generally perceived as the best performing method. All entries in the tables that follow are percentages.

Table 1 presents results for $n = 35$, $p = 2$ and different values for k. We vary k to analyze the effect of the number of nuisance parameters on the different tests. At the outset, we note that the original likelihood ratio test is considerably oversized, especially as the number of nuisance parameters increases. For instance, when $k = 7$ and $\alpha = 5\%$, the null rejection rate exceeds 16%. The tendency of the test to overreject is attenuated by the Bartlett correction, the resulting Bartlett-adjusted test displaying smaller size distortions. For example, its null rejection rate for the same situation was

Table 1
Null rejection rates, $n = 35$ and $p = 2$

k	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$				$\alpha = 0.5\%$			
	LR	LR*	LR$_m$	LR$_m^*$	LR	LR*	LR$_m$	LR$_m^*$	LR	LR*	LR$_m$	LR$_m^*$	LR	LR*	LR$_m$	LR$_m^*$
2	13.4	10.6	8.9	9.8	7.5	5.4	4.3	4.9	1.6	1.1	0.7	0.9	0.9	0.5	0.4	0.5
3	14.5	10.9	8.7	9.9	8.2	5.9	4.2	4.8	2.3	1.2	0.8	1.0	1.2	0.7	0.3	0.4
4	16.3	1.2	8.3	9.2	9.5	5.7	4.1	4.7	2.7	1.3	0.8	1.0	1.7	0.8	0.4	0.5
5	19.2	12.8	8.7	9.6	1.9	7.2	4.3	4.8	3.8	1.6	0.7	0.9	2.4	0.9	0.4	0.4
6	22.3	14.4	8.1	9.1	14.6	7.8	4.0	4.5	5.2	2.1	0.7	1.0	3.3	1.3	0.4	0.4
7	24.9	15.6	8.9	9.6	16.6	9.0	4.2	4.6	6.3	2.5	0.8	1.0	4.1	1.5	0.4	0.4
8	30.4	18.9	8.4	9.3	2.8	1.7	3.9	4.3	9.7	3.7	0.8	0.9	6.9	2.3	0.3	0.4

Table 2
Null rejection rates, $n=35$ and $k=5$

p	$\alpha=10\%$				$\alpha=5\%$				$\alpha=1\%$				$\alpha=0.5\%$			
	LR	LR*	LR_m	LR_m^*	LR	LR*	LR_m	LR_m^*	LR	LR*	LR_m	LR_m^*	LR	LR*	LR_m	LR_m^*
1	16.3	1.6	9.1	9.7	9.6	6.1	4.4	4.7	3.0	1.6	0.9	1.1	1.8	0.9	0.4	0.4
2	19.2	12.8	8.7	9.6	1.9	7.2	4.3	4.8	3.8	1.6	0.7	0.9	2.4	0.9	0.4	0.4
3	2.8	13.9	8.8	9.7	14.0	7.5	4.2	4.8	4.7	1.9	0.8	1.0	2.9	1.0	0.4	0.5
4	26.6	15.8	8.7	9.9	17.2	9.1	4.1	4.8	6.3	2.4	0.9	1.0	4.1	1.4	0.5	0.5
5	29.6	17.3	9.0	10.0	20.0	9.9	4.5	5.0	8.0	2.8	0.8	1.0	5.3	1.7	0.4	0.5
6	36.0	2.2	9.1	10.3	25.9	13.0	4.7	5.4	1.2	4.4	1.1	1.3	8.0	2.8	0.6	0.8
7	46.1	27.6	9.4	10.4	34.9	18.0	4.4	5.1	17.3	6.9	0.8	1.0	12.9	4.2	0.4	0.4
8	49.4	28.8	9.2	10.1	37.4	19.2	4.2	4.7	18.9	7.2	0.9	1.0	14.1	4.7	0.3	0.4

9.0%. The modified profile likelihood ratio test tends to overcorrect the liberal tendency of the original test, displaying null rejection rates that are smaller than the nominal level of the test. Again, when $k=7$ and $\alpha=5\%$, its null rejection rate was 4.2%. The Bartlett correction applied to this test brings its empirical type I error probability closer to the nominal size of the test. For the situation singled out above, its null rejection rate was 4.6%. The above conclusions also hold for other nominal levels, even for very small ones.

Table 2 contains results for the situation where $n=35$, $k=5$ and $p=1,\dots,8$. Again, the results show that the original likelihood ratio test is considerably oversized, the more so the larger the number of covariates used in the specification of the skedastic function; e.g., when $p=5$ and $\alpha=5\%$, the rejection rate of the test under the null hypothesis was 20.0%, i.e., four times the nominal level selected for the test. The other conclusions drawn from Table 1 also hold here. For the same situation ($p=5$ and $\alpha=5\%$), the null rejection rates of the Bartlett-corrected likelihood ratio test, the modified profile likelihood ratio test, and its Bartlett-adjusted version were 9.9%, 4.5% and 5.0%, respectively.

In Table 3, we fix the value of p at 2 and vary the sample size: $n=30,50,100$ for $k=5,8$. As the sample increases, the null rejection rates of all four tests become closer to the nominal level of the test, as expected. We also note that even for $n=100$ the original likelihood ratio test is still oversized. When $k=8$, $\alpha=5\%$ and $n=50$, the null rejection rates of the likelihood ratio test, its Bartlett-corrected version, the modified profile likelihood ratio test, and its Bartlett-adjusted version were, respectively, 13.6%, 7.7%, 4.6% and 5.0%.

The test of equality of variances for the one factor and three levels model is considered in Table 4. Here, there are three normal populations with means μ_1,μ_2,μ_3 and variances $\sigma_1^2,\sigma_2^2,\sigma_3^2$, and the null hypothesis of interest is $\mathscr{H}_0$: $\sigma_1^2=\sigma_2^2=\sigma_3^2$. Note that $p=2$ and $k=3$. The number of observations from each of the three populations was n_1,n_2,n_3 ($n_1+n_2+n_3=n$). The simulation was performed using $n_1=n_2=n_3=n/3$. The figures in Table 4 yield conclusions that are in agreement with the ones drawn from Tables 1–3.

Table 3
Null rejection rates, $p=2$ and several sample sizes

n	α	$k=5$				$k=8$			
		LR	LR*	LR_m	LR_m^*	LR	LR*	LR_m	LR_m^*
30	10.0	25.0	16.4	9.1	10.2	36.6	23.4	8.3	9.2
	5.0	16.7	9.9	4.2	5.0	27.5	14.9	3.7	4.3
	1.0	6.6	2.8	0.8	1.0	13.4	4.8	0.6	0.8
	0.5	4.2	1.5	0.4	0.6	9.5	3.1	0.3	0.4
50	10.0	16.2	11.2	9.2	9.8	21.6	13.8	9.4	10.1
	5.0	9.2	5.8	4.3	4.7	13.6	7.7	4.6	5.0
	1.0	2.5	1.3	0.9	1.0	4.7	2.1	1.0	1.1
	0.5	1.5	0.7	0.4	0.5	3.0	1.1	0.5	0.4
100	10.0	12.0	9.8	9.1	9.4	14.5	10.8	9.8	10.0
	5.0	6.2	5.0	4.6	4.9	8.1	5.7	4.9	5.2
	1.0	1.6	1.0	0.9	1.0	2.2	1.3	1.0	1.1
	0.5	0.8	0.5	0.4	0.4	1.3	0.6	0.5	0.6

Table 4
Null rejection rates, model with one factor and three levels, $p=2$ and $k=3$

n	α	LR	LR*	LR_m	LR_m^*
9	10.0	28.2	14.4	5.4	6.8
	5.0	18.6	8.1	2.4	3.2
	1.0	7.4	2.2	0.3	0.5
	0.5	5.0	1.2	0.1	0.2
18	10.0	16.6	10.6	8.5	9.2
	5.0	9.6	5.8	4.2	4.7
	1.0	3.1	1.3	0.8	0.9
	0.5	1.9	0.7	0.4	0.5
27	10.0	14.6	10.6	9.3	9.9
	5.0	8.2	5.5	4.7	5.0
	1.0	2.1	1.1	0.9	1.0
	0.5	1.2	0.6	0.4	0.5
36	10.0	12.8	10.1	9.4	9.8
	5.0	7.2	5.3	4.7	5.0
	1.0	1.8	1.1	0.9	1.1
	0.5	1.0	0.5	0.4	0.5
45	10.0	12.3	10.1	9.3	9.7
	5.0	6.5	5.2	4.7	5.0
	1.0	1.5	1.1	1.0	1.1
	0.5	0.9	0.5	0.5	0.5

Table 5
Nonnull rejection rates, $n = 30,\ p = 2$

δ	$k = 5$		$k = 8$	
	LR_m	LR_m^*	LR_m	LR_m^*
0.5	11.1	12.3	11.2	12.6
1.0	26.3	28.2	25.0	26.9
1.5	45.2	47.2	48.7	51.0
2.0	63.8	65.9	70.1	72.2
2.5	73.9	75.5	75.7	77.3
3.0	91.6	92.3	83.3	85.6
3.5	99.9	99.9	92.2	92.9

Table 5 collects results obtained from simulations carried out under the alternative hypothesis (heteroskedasticity) for $n = 30,\ p = 2$ and different values of $\delta_1 = \delta_2 = \delta$ at the 10% nominal level. It is noteworthy that these power simulations correspond to the setting in Table 3 for $n = 30$. We only compare the power of the modified profile likelihood ratio test and its Bartlett-corrected variant, since the remaining two tests are considerably oversized and cannot be recommended. The results indicate that there is no loss in power derived from using the Bartlett adjustment proposed in this paper. The power of the two tests are similar, with the slight advantage of the corrected test steming from its smaller size distortion.

Since the results for the Bartlett-corrected modified profile likelihood ratio test presented above are obtained by correcting the exact form of the test and Ferrari and Cribari-Neto (2002) have corrected the test proposed by Simonoff and Tsai (1994), which is in fact an approximation to the exact form, it is interesting to see what happens when one corrects this approximated test instead of the modified profile likelihood ratio test in its original form. We note that the modified profile likelihood ratio test and its Bartlett-corrected display more reliable behavior when the approximation proposed by Simonoff and Tsai (1994) is not used. For instance, consider the simulation results in Table 3 with $k = 8$, $\alpha = 5\%$ and $n = 50$. As noted earlier, the rejection rates for the modified profile likelihood ratio test and its Bartlett-adjusted version are 4.6% and 5.0%, respectively. The corresponding rejection rates for the two tests that employ the approximation used by Simonoff and Tsai (1994) are 3.7% and 4.1%.

It is well known that the likelihood ratio test is extremely oversized under nonnormality. We have obtained some additional simulation results by generating the errors from some nonnormal distributions. We noticed that, even under slight departures from normality (e.g., when the u_t's are obtained independently from t_{20} distributions), all tests are strongly oversized. However, the modified tests display much smaller size distortions than the uncorrected likelihood ratio test.

Overall, the likelihood ratio test can be severely liberal, overrejecting the null hypothesis more often than expected based on the selected nominal level for the test. The Bartlett correction to the likelihood ratio test does bring the empirical null rejection rate closer to the nominal level of the test, but it does not fully correct its liberal tendency. The modified profile likelihood ratio test introduces, on the other hand, an overcorrection, thus delivering an undersized test. The Bartlett adjustment to this test brings the size distortions closer to zero; it yields the most accurate of the four tests.

5. Applications

At the outset, we shall consider the data analyzed by Montgomery et al. (2001, p. 76). The application involves a soft drink bottler who wishes to predict the amount of time required to service the vending machines in an outlet. The service activity includes stocking the machines with cans and bottles of beverages and performing minor maintenance or housekeeping. The response (y) is the time (in minutes) spent on servicing soft drinks machines, and the covariates are the number of beverage cases stocked (x_2) and the distance (in feet) traveled (x_3). The model used is $y_l=\beta_1+\beta_2 x_{l2}+\beta_3 x_{l3}+u_l$, $l=1,\ldots,25$. We assume that $u_l \sim \mathcal{N}(0,\sigma^2\exp\{\delta_1 x_{l2}+\delta_2 x_{l3}\})$ and $\mathrm{cov}(u_l,u_m)=0$ for all $l \neq m$. Standard diagnostic analyses suggest that observations 9 and 22 are leverage points, the former also classifying as an influential observation; see Montgomery et al. (2001, pp. 210, 213, 215, 216, 217). In what follows, we shall work with the incomplete data set, i.e., we shall exclude observations 9 and 22 from the data. The maximum likelihood parameter estimates are $\hat\beta_1=4.643$, $\hat\beta_2=1.456$ and $\hat\beta_3=0.011$. The estimate of σ^2 is $\hat\sigma^2=5.360$ and the coefficient of determination, R^2, equals 0.9072. These estimates were obtained imposing $\delta_1=\delta_2=0$. Diagnostic analyses for the incomplete data set do not show evidence of heteroskedasticity. Our main interest lies in testing $\mathscr{H}_0$: $\delta_1=\delta_2=0$ (homoskedasticity) against a two-sided alternative (heteroskedasticity). Rejection of the null hypothesis would suggest that the nonconstant response variance should be modeled as well. The test statistics are $\mathrm{LR}=4.825$, $\mathrm{LR}^*=3.717$, $\mathrm{LR}_m=4.127$ and $\mathrm{LR}_m^*=4.352$; the respective p-values are 0.090, 0.156, 0.127 and 0.113. (The computer code for computing these statistics using the `Ox` matrix programming language and the data set are available at `http://www.de.ufpe.br/~cribari/FerrariCysneirosCribari.zip`.) The standard likelihood ratio test yields rejection of the null hypothesis (thus suggesting the presence of heteroskedasticity) at the 10% nominal level whereas the null of homoskedasticity is not rejected, at the same nominal level, when the remaining tests (i.e., those based on LR^*, LR_m and LR_m^*) are used. That is, the Bartlett-corrected, the adjusted profile likelihood ratio and the corrected adjusted profile likelihood ratio tests lead to inference different from that reached by the standard likelihood ratio test.

We now move to the data set presented in Simonoff and Tsai (1994, Table 1). The data represent the monthly excess returns over the riskless rate for the market (x) and for the Acme Cleveland Corporation (y) from January 1986 through December 1990. The model used is $y_l=\beta_1+\beta_2 x_l+u_l$, $l=1,\ldots,59$. We assume that $u_l \sim \mathcal{N}(0,\sigma^2\exp\{\delta x_l\})$ and $\mathrm{cov}(u_l,u_m)=0$ for all $l \neq m$. The test statistics for testing $\mathscr{H}_0$: $\delta=0$ (homoskedasticity) against a two-sided alternative (heteroskedasticity) are $\mathrm{LR}=3.329$, $\mathrm{LR}^*=3.119$, $\mathrm{LR}_m=2.968$ and $\mathrm{LR}_m^*=3.071$. The respective p-values are 0.068, 0.077, 0.085 and 0.080, all the tests indicating marginal heteroskedasticity. Here, the sample size is large ($n=59$) relative to the number of covariates in X ($k=2$) and to the number of covariates in the variance specification ($p = 1$). Hence, as expected, the finite sample corrections to the likelihood ratio statistic do not have significant impact.

6. Concluding remarks

Practitioners commonly test for the presence of heteroskedasticity when estimating linear regression structures in order to decide what is the appropriate modeling strategy. It is thus important to devise reliable tests for detecting nonconstant conditional variances. In this paper we have derived a Bartlett correction that can be applied to likelihood ratio statistics obtained using modified profile likelihood. The results allow for skedastic functions that involve more than one parameter, and thus generalize results available in the literature. The numerical results presented compared the finite-sample performance of four tests, namely: the original likelihood ratio test, its Bartlett-corrected version, the modified profile likelihood ratio test, and the Bartlett-adjusted modified profile likelihood ratio test we proposed. Overall, the numerical results favor the latter, i.e., they favor the test obtained from applying a Bartlett correction to the statistic of the modified profile likelihood ratio test. In short, the modification of the profile likelihood attenuates the effects of nuisance parameters on the inference, and the Bartlett correction yields a faster convergence rate of the test statistic to its first-order asymptotic distribution. We, therefore, encourage practitioners to use the test proposed in the present paper in applications.

Acknowledgements

We gratefully acknowledge partial financial support from CAPES, CNPq, FAPESP and FINEP.

References

Barndorff-Nielsen, O.E., Hall, P., 1988. On the level-error after Bartlett adjustment of the likelihood ratio statistic. Biometrika 75, 374–378.

Cordeiro, G.M., 1993. Bartlett corrections and bias correction for two heteroscedastic regression models. Comm. Statist. Theory Methods 22, 169–188.

Cox, D.R., Reid, N., 1987. Parameter orthogonality and approximate conditional inference. J. Roy. Statist. Soc. B 49, 1–39.

Cribari-Neto, F., Cordeiro, G.M., 1996. On Bartlett and Bartlett-type corrections. Econom. Rev. 15, 339–367.

DiCiccio, T.J., Stern, S.E., 1994. Frequentist and Bayesian Bartlett correction of test statistics based on adjusted profile likelihoods. J. Roy. Statist. Soc. B 56, 397–408.

Doornik, J.A., 2001. Ox: an Object-oriented Matrix Programming Language, 4th Edition. Timberlake Consultants and Oxford, London, http://www.nuff.ox.ac.uk/Users/Doornik/.

Ferrari, S.L.P., Cribari-Neto, F., 2002. Corrected modified profile likelihood heteroskedasticity tests. Statist. Probab. Lett. 57, 353–361.

Lawley, D.N., 1956. A general method for approximating to the distribution of the likelihood ratio criteria. Biometrika 71, 233–244.

Montgomery, D.C., Peck, E.A., Vining, G.G., 2001. Introduction to Linear Regression Analysis, 3rd Edition. Wiley, New York.

Simonoff, J.S., Tsai, C.-H., 1994. Use of modified profile likelihood for improved tests of constancy of variance in regression. Appl. Statist. 43, 357–370.