Brazilian Journal of Probability and Statistics
2014, Vol. 28, No. 1, 83–95
DOI: 10.1214/12-BJPS196

New heteroskedasticity-robust standard errors for the linear regression model

Francisco Cribari-Neto and Maria da Glória A. Lima

Universidade Federal de Pernambuco

Abstract. Linear regressions fitted to cross-sectional data oftentimes display heteroskedasticity, that is, error variances that are not constant. A common modeling strategy consists of estimating the regression parameters by ordinary least squares and then performing hypothesis testing inference using standard errors that are robust to heteroskedasticity. These tests have the correct size asymptotically regardless of whether the error variances are constant. In finite samples, however, they can be quite size-distorted. In this paper, we propose new heteroskedasticity-consistent covariance matrix estimators that deliver more reliable testing inferences in samples of small sizes.

1 Introduction

The linear regression model is commonly used by practioners in empirical analyses in fields such as chemistry, economics, finance, medicine, physics, among others. An assumption that is frequently violated is that of homoskedasticity, that is, the assumption that all errors share the same variance. Standard hypothesis tests (e.g., the usual t test) do not have the correct size when the errors are heteroskedastic, nor even asymptotically, since they are based on an inconsistent covariance matrix estimator. Several heteroskedasticity-robust standard errors were proposed in the literature, the most commonly employed being that obtained from White's (1980) covariance matrix estimator. Halbert White proposed an estimator for the covariance matrix of the vector of regression parameters least squares estimator that is consistent under both homoskedasticity and heteroskedasticity of unknown form. Using his estimator, one can perform hypothesis testing inference that is asymptotically correct in the sense that the type I error probability will approach the selected nominal size as the sample size grows. Nonetheless, White's (1980) estimator, which we shall denote by HC0, is typically quite biased and can deliver unreliable testing inference in small to moderately large samples, more so under leveraged data. Several alternatives were proposed in the literature, such as the HC1 (Hinkley, 1977), HC2 (Horn, Horn and Duncan, 1975), HC3 (Davidson and MacKinnon, 1993), HC4 (Cribari-Neto, 2004) and HC5 (Cribari-Neto, Souza

Key words and phrases. Covariance matrix estimation, heteroskedasticity, linear regression, quasi-t test.

Received February 2012; accepted May 2012.

and Vasconcellos, 2007) estimators. These alternative estimators include finite-sample corrections and their associated tests are typically less size-distorted than the HC0-based test. Bias-corrected variants of White's estimator were obtained by Cribari-Neto, Ferrari and Cordeiro (2000) and by Qian and Wang (2001); see also Cribari-Neto and Lima (2010).

Under heteroskedasticity, the ordinary least squares estimator of the vector of regression parameters is no longer efficient in the class of linear and unbiased estimators, but it remains unbiased, consistent and asymptotically normal. A t test in which the usual least squares standard error in the test statistic is replaced by a heteroskedasticity-robust standard error is said to be a "quasi-t" test. Since the variance estimator is consistent, the test statistic null distribution converges to the standard Gaussian distribution $\mathcal{N}(0, 1)$. The test can then be based on standard normal (asymptotic) critical values. These critical values are used as an approximation to the (unknown) exact critical values. As a result, size distortions are likely to occur in finite samples. Such distortions are typically more pronounced when the data contain high leverage observations (Chesher and Jewitt, 1987). In particular, the HC0 (White) estimator tends to be "optimistic" (positively biased) and the HC0-based test tends to be liberal (anticonsertative). Long and Ervin (2000) recommend the use of the HC3-based test whereas the numerical evidence in Cribari-Neto (2004) favors HC4-based testing inference.

The chief purpose of this paper is to propose alternative heteroskedasticity-consistent covariance matrix estimators. As shown by our numerical evaluations, the proposed estimators deliver more reliable finite sample testing inference in small samples under both homoskedasticity and unequal error variances. The proposed estimators are based on a class of consistent estimators given in Cribari-Neto and Lima (2010). They are sandwich-like estimators: a slice of bread (involving the matrix of covariates), meat (involving squared regression residuals, which are taken as estimators for the underlying unknown variances) and another slice of bread (the transpose of the first slice of bread). Unlike the estimators of Cribari-Neto and Lima (2010), our estimators include finite sample correction factors in both the "bread" and the "meat." As a result, the proposed estimators yield more reliable testing inference in small samples and under leveraged data in the sense that hypothesis tests whose statistics employ standard errors obtained from them have small size distortions in small samples, even when the data contain high leveraged observations. The results of numerical evaluations we present clearly favor hypothesis testing inference based on two estimators we propose. Inferences drawn from tests whose statistics employ standard errors from such estimators are considerably more reliable than those drawn from tests based on competing estimators, including tests based on the estimators of Cribari-Neto and Lima (2010). For example, in the first numerical evaluation discussed Section 4 the null rejection rate of the test based on White's estimator (HC0) at the 5% nominal level under strong heteroskedasticity exceeds 40% (eight times the nominal level!) whereas our best performing test displays nearly correct size (4%). This is a substantial improvement.

The paper unfolds as follows. Section 2 introduces the regression model and some covariance matrix estimators that are consistent under both homoskedasticity and heteroskedasticity of unkown form. New heteroskedasticity-consistent covariance matrix estimators are proposed in Section 3. Section 4 contains the results of several numerical evaluations. As we shall see, the numerical evidence favors hypothesis testing inference based on the estimators proposed in this paper. Such tests display better control of the type I error probability than competing tests (e.g., HC0, HC3 and HC4 tests and also the tests based on the estimators proposed in Furno, 1996 and Cribari-Neto and Lima, 2010). Finally, Section 5 offers some concluding remarks.

2 The model and heteroskedasticity-robust standard errors

The linear regression model is written as $y = X\beta + \varepsilon$, where y and ε are n-vectors of responses and errors, respectively, X is an $n \times p$ matrix of fixed regressors ($\operatorname{rank}(X) = p < n$) and $\beta = (\beta_1, \dots, \beta_p)'$ is a p-vector of unknown regression parameters, n being the sample size, that is, the number of observations in the sample. Each error ε_i has mean zero, variance $0 < \sigma_i^2 < \infty$, $i = 1, \dots, n$, and is uncorrelated with ε_j whenever $j \neq i$. The errors covariance matrix is thus $\Omega = \operatorname{cov}(\varepsilon) = \operatorname{diag}\{\sigma_1^2, \dots, \sigma_n^2\}$.

Estimation of β can be carried out by ordinary least squares (OLS), the resulting estimator being $\widehat{\beta} = (X'X)^{-1}X'y$. Its covariance matrix is $\Psi = \operatorname{cov}(\widehat{\beta}) = P\Omega P'$, where $P = (X'X)^{-1}X'$. Under homoskedasticity, $\sigma_i^2 = \sigma^2$, $i = 1, \dots, n$, where $\sigma^2 > 0$, and hence $\Psi = \sigma^2(X'X)^{-1}$. The covariance matrix Ψ can then be easily estimated as $\widehat{\Psi} = \widehat{\sigma}^2(X'X)^{-1}$, where $\widehat{\sigma}^2 = (y - X\widehat{\beta})'(y - X\widehat{\beta})/(n - p) = \widehat{\varepsilon}'\widehat{\varepsilon}/(n - p)$. The vector of least squares residuals is given by $\widehat{\varepsilon} = (\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_n)' = (I - H)y$, where $H = X(X'X)^{-1}X' = XP$ and I is the $n \times n$ identity matrix. (H is known as "the hat matrix," since the Hy yields $\widehat{y}$, the vector of fitted values.)

When the errors are heteroskedastic, one can use the OLSE of β coupled with a consistent covariance matrix estimator in order to perform hypothesis testing inference. In an influential paper, White (1980) obtained a consistent estimator for Ψ. His estimator is consistent under both homoskedasticity and heteroskedasticity of unknown form, and can be written as

$$\mathrm{HC0} = \widehat{\Psi}_0 = P\widehat{\Omega}P',$$

where $\widehat{\Omega} = \operatorname{diag}\{\widehat{\varepsilon}_1^2, \dots, \widehat{\varepsilon}_n^2\}$. Using HC0, one can then obtain standard errors for the parameter estimates that are robust against heteroskedasticity. A noteworthy shortcoming of HC0 lies in its finite-sample behavior: HC0 is typically quite biased, especially when the data contain high leverage points; see, for example, Chesher and Jewitt (1987). In particular, HC0 tends to understimate the true variances and, as a result, associated hypothesis tests tend to be considerably liberal when the sample size is small.

Several variants of White's estimators were proposed in the literature; for instance, HCs $= \widehat{\Psi}_s = P\widehat{\Omega}_s P'$, for $s = 1, \ldots, 4$, where $\widehat{\Omega}_s$ is a diagonal matrix of the form $\{\widehat{\varepsilon}_1^2/\delta_1, \ldots, \widehat{\varepsilon}_n^2/\delta_n\}$. For HC1, $\delta_i = (n-p)/n$; for HC2, $\delta_i = 1 - h_i$; for HC3, $\delta_i = (1-h_i)^2$ and for HC4, $\delta_i = (1-h_i)^{\gamma_i}$, with $\gamma_i = \min\{4, nh_i/p\}$ $(i = 1, \ldots, n)$. Here, h_i denotes the ith diagonal element of H.

A sequence of bias-corrected HC0 estimators was derived by Cribari-Neto, Ferrari and Cordeiro (2000). They used an iterative bias correction mechanism. Their chain of estimators was obtained by correcting HC0, then correcting the resulting adjusted estimator, and so on. Let $(A)_d$ denote the diagonal matrix obtained by setting the nondiagonal elements of the square matrix A equal to zero and consider the following recursive function of an $n \times n$ diagonal matrix A:

$$M^{(k+1)}(A) = M^{(1)}\big(M^{(k)}(A)\big), \qquad k = 0, 1, \ldots,$$

where $M^{(0)}(A) = A$, $M^{(1)}(A) = \{HA(H-2I)\}_d$. It can be shown that

$$\mathbb{E}(\widehat{\Omega}) = \big\{(I-H)\Omega(I-H)\big\}_d = M^{(1)}(\Omega) + \Omega. \tag{2.1}$$

Therefore, the biases of $\widehat{\Omega}$ and $\widehat{\Psi}$ as estimators of Ω and Ψ are

$$B_{\widehat{\Omega}}(\Omega) = \mathbb{E}(\widehat{\Omega}) - \Omega = \big\{H\Omega(H-2I)\big\}_d = M^{(1)}(\Omega)$$

and

$$B_{\widehat{\Psi}}(\Omega) = \mathbb{E}(\widehat{\Psi}) - \Psi = PB_{\widehat{\Omega}}(\Omega)P'.$$

Cribari-Neto, Ferrari and Cordeiro (2000) defined the following bias-corrected estimator:

$$\widehat{\Omega}^{(1)} = \widehat{\Omega} - B_{\widehat{\Omega}}(\widehat{\Omega}).$$

This estimator can be adjusted for bias as well: $\widehat{\Omega}^{(2)} = \widehat{\Omega}^{(1)} - B_{\widehat{\Omega}^{(1)}}(\widehat{\Omega})$. After k iterations of the bias-correcting scheme one obtains $\widehat{\Omega}^{(k)} = \widehat{\Omega}^{(k-1)} - B_{\widehat{\Omega}^{(k-1)}}(\widehat{\Omega})$.

The authors have shown that the kth order bias-corrected estimator and its bias can be written as $\widehat{\Omega}^{(k)} = \sum_{j=0}^{k}(-1)^j M^{(j)}(\widehat{\Omega})$ and $B_{\widehat{\Omega}^{(k)}}(\Omega) = (-1)^k M^{(k+1)}(\Omega)$, for $k = 1, 2, \ldots$. They then defined a sequence of bias-corrected covariance matrix estimators as $\{\widehat{\Psi}^{(k)}, k = 1, 2, \ldots\}$, where $\widehat{\Psi}^{(k)} = P\widehat{\Omega}^{(k)}P'$. The bias of $\widehat{\Psi}^{(k)}$ is $B_{\widehat{\Psi}^{(k)}}(\Omega) = (-1)^k PM^{(k+1)}(\Omega)P'$, $k = 1, 2, \ldots$.

They have shown also that the bias of HC0 is $B_{\widehat{\Psi}}(\Omega) = PB_{\widehat{\Omega}}(\Omega)P' = O(n^{-2})$ and $B_{\widehat{\Psi}^{(k)}}(\Omega) = O(n^{-(k+2)})$. That is, the bias of the kth corrected estimator is of order $O(n^{-(k+2)})$, whereas the bias of Halbert White's estimator is $O(n^{-2})$.

Yet another alternative estimator was proposed by Qian and Wang (2001). Let $K = (H)_d = \mathrm{diag}\{h_1, \ldots, h_n\}$ and let $C_i = X(X'X)^{-1}x_i'$ denote the ith column of H, x_i being the ith row of X. Define

$$D^{(1)} = \mathrm{diag}\{d_i\} = \mathrm{diag}\big\{\big(\widehat{\varepsilon}_i^2 - \widehat{b}_i\big)g_{ii}\big\},$$

where $g_{ii} = (1 + C_i' K C_i - 2h_i^2)^{-1}$ and $\widehat{b}_i = C_i'(\widehat{\Omega} - 2\widehat{\varepsilon}_i^2 I)C_i$. The Qian–Wang estimator can be written as

$$\widehat{V}^{(1)} = P D^{(1)} P'.$$

Cribari-Neto and Lima (2010) have shown that

$$D^{(1)} = [\widehat{\Omega} - \{H\widehat{\Omega}(H - 2I)\}_d]G,$$

where

$$G = \{I + HKH - 2KK\}_d^{-1}.$$

They then derived a sequence of bias-corrected covariance matrix estimators that starts at $\widehat{V}^{(1)}$.

An estimator that does not employ OLS residuals was proposed by Furno (1996). It is given by

$$\mathrm{HC0}_R = (X'WX)^{-1} X'W\widehat{\Psi}_R WX(X'WX)^{-1},$$

where W is an $n \times n$ diagonal matrix whose ith diagonal element is $w_i = \min(1, c/h_i)$. She suggests using $c = 1.5p/n$. Here,

$$\widehat{\Psi}_R = \mathrm{diag}\{\widetilde{e}_1^2, \ldots, \widetilde{e}_n^2\}, \tag{2.2}$$

where $\widetilde{e}_i$ is the ith weighted least squares residual. A robustified HC3-like estimator ($\mathrm{HC3}_R$) is obtained by replacing $\widehat{\Psi}_R$ in (2.2) by $\widehat{\Psi}_{3R} = \mathrm{diag}\{\widetilde{e}_1/(1 - h_1^*)^2, \ldots, \widetilde{e}_n/(1 - h_n^*)^2\}$, where h_i^* is the ith diagonal element of $\sqrt{W}X(X'WX)^{-1}X'\sqrt{W}$.

3 New covariance matrix estimators

In what follows, we shall propose new heteroskedasticity-consistent covariance matrix estimators. At the outset, note that Halbert White's HC0 estimator can be written as

$$\mathrm{HC0} = \widehat{\Psi}_0 = P\widehat{\Omega}_0 P' = PD_0\widehat{\Omega}P',$$

where $D_0 = I$. Additionally, note that

(i) $\mathrm{HC1} = \widehat{\Psi}_1 = P\widehat{\Omega}_1 P' = PD_1\widehat{\Omega}P'$, $D_1 = (n/(n-p))I$;
(ii) $\mathrm{HC2} = \widehat{\Psi}_2 = P\widehat{\Omega}_2 P' = PD_2\widehat{\Omega}P'$, $D_2 = \mathrm{diag}\{1/(1-h_i)\}$;
(iii) $\mathrm{HC3} = \widehat{\Psi}_3 = P\widehat{\Omega}_3 P' = PD_3\widehat{\Omega}P'$, $D_3 = \mathrm{diag}\{1/(1-h_i)^2\}$;
(iv) $\mathrm{HC4} = \widehat{\Psi}_4 = P\widehat{\Omega}_4 P' = PD_4\widehat{\Omega}P'$, $D_4 = \mathrm{diag}\{1/(1-h_i)^{\gamma_i}\}$ and $\gamma_i = \min\{4, nh_i/p\}$.

In what follows, we shall denote these estimators by HCs, $s = 0, 1, 2, 3, 4$.

We have seen in (2.1) that $\mathbb{E}(\widehat{\Omega}) = M^{(1)}(\Omega) + \Omega$. Then,

$$\mathbb{E}(\widehat{\Omega}_s) = \mathbb{E}(D_s\widehat{\Omega}) = D_s\mathbb{E}(\widehat{\Omega}) = D_sM^{(1)}(\Omega) + D_s\Omega$$

and

$$B_{\widehat{\Omega}_s}(\Omega) = \mathbb{E}(\widehat{\Omega}_s) - \Omega = D_sM^{(1)}(\Omega) + (D_s - I)\Omega.$$

Cribari-Neto and Lima (2010) introduced the following estimator:

$$\widehat{\Psi}_s^{(1)} = P\widehat{\Omega}_s^{(1)}P',$$

where $\widehat{\Omega}_s^{(1)} = \widehat{\Omega}_s - B_{\widehat{\Omega}_s}(\widehat{\Omega}) = \widehat{\Omega} - D_sM^{(1)}(\widehat{\Omega})$. Thus,

$$\begin{aligned}\mathbb{E}\big(\widehat{\Omega}_s^{(1)}\big) &= \mathbb{E}(\widehat{\Omega}) - D_sM^{(1)}\big(\mathbb{E}(\widehat{\Omega})\big)\\ &= M^{(1)}(\Omega) - D_sM^{(1)}(\Omega) + \Omega - D_sM^{(2)}(\Omega).\end{aligned}$$

Hence, under homoskedasticity, $\mathbb{E}(\widehat{\Omega}_s^{(1)}) = \sigma^2 A_s$, where $A_s = (I - K) + D_s\{K + HKH - 2KK\}_d$.

It then follows that the expected value of $\widehat{\Psi}_s^{(1)}$ when $\Omega = \sigma^2 I$ (homoskedastic errors) is given by $\mathbb{E}(\widehat{\Psi}_s^{(1)}) = \mathbb{E}(P\widehat{\Omega}_s^{(1)}P') = \sigma^2 PA_sP'$. They then proposed a second estimator:

$$\widehat{\Psi}_{sA} = P\widehat{\Omega}_s^{(1)}A_s^{-1}P', \tag{3.1}$$

which is unbiased under equal error variances: $\mathbb{E}(\widehat{\Psi}_{sA}) = \mathbb{E}(P\widehat{\Omega}_s^{(1)}A_s^{-1}P') = \Psi$.

It is noteworthy that the Qian-Wang estimator introduced in Section 2 is a particular case of $\widehat{\Psi}_{sA}$; it is obtained by setting $s = 0$, that is, when $D_0 = I$. Note that $A_0 = G^{-1}$.

We shall now introduce two new classes of heterokedasticity-consistent covariance matrix estimators. They are based on the estimators proposed by Cribari-Neto and Lima (2010). Our estimators use the "meat" of such estimators, but employ different slices of "bread." The motivation behind them is that the "bread" should also include finite sample corrections based on the different leverage measures, as it is done in the "meat." The corrections we employ are different from that used by Furno (1996); they are similar to the corrections used in "meat" of the HC2, HC3 and HC4 estimators. We also note that, unlike Furno's (1996) estimator, the estimators we propose use standard OLS residuals.

Our first class of estimators is given by

$$\widehat{\Psi}_s^{\delta} = P_\delta D_s\widehat{\Omega}P_\delta' = P_\delta\widehat{\Omega}_sP_\delta', \tag{3.2}$$

where D_s is as in the definition of HCs, $s = 0, \ldots, 4$, $P_\delta = (X'D^\delta X)^{-1}X'$ and $D^\delta = \mathrm{diag}\{(1-h_1)^\delta, \ldots, (1-h_n)^\delta\}$, $0 \le \delta \le 1$. Our second class of estimators is

$$\widehat{\Psi}_{sA}^{\delta} = P_\delta\widehat{\Omega}_s^{(1)}A_s^{-1}P_\delta'. \tag{3.3}$$

It is defined using the “meat” of estimator $\widehat{\Psi}_{sA}$, which is given in (3.1). It is noteworthy that none of our estimators is unbiased under homoskedasticity, unlike the estimator in (3.1), since we modified the bread part of the estimator by making it include a finite sample correction factor. The well known HC*s* estimators ($s = 0, \ldots, 4$) are special cases of (3.2); these estimators are obtained by setting $\delta = 0$. Likewise, the estimators in (3.1) proposed by Cribari-Neto and Lima (2010) are obtained as special cases of the estimators given in (3.3) when $\delta = 0$. As we shall see in the next section, values of δ between 0.5 and 1.0 usually yield reliable inferences in small samples.

Heteroskedasticity-consistent covariance matrix estimators are typically used to perform testing inferences on the parameters that index the regression model. As noted earlier, heteroskedasticity-robust standard errors (square roots of the diagonal elements) are used in quasi-t test statistics. The resulting tests have the correct size asymptotically regardless of whether or not the error variances are constant. The results of numerical evaluations that compare the finite sample performances of different quasi-t tests are presented in the next section.

4 Numerical results

We shall now present Monte Carlo simulation results on the finite-sample performance of hypothesis tests based on different heteroskedasticity-consistent standard errors. We shall consider quasi-t tests whose statistics employ standard errors obtained from the following estimators: OLS (i.e., $\widehat{\sigma}^2(X'X)^{-1}$), HC0 (White, 1980), HC0$_R$ (Furno, 1996), HC3 (Davidson and MacKinnon, 1993), $\widehat{\Psi}_3^{0.5}$, HC4 (Cribari-Neto, 2004), $\widehat{\Psi}_4^{0.5}$, $\widehat{\Psi}_{4A}$ (Cribari-Neto and Lima, 2010), $\widehat{\Psi}_{3A}$ (Cribari-Neto and Lima, 2010), $\widehat{\Psi}_{4A}^{0.5}$, $\widehat{\Psi}_{3A}^{0.5}$, $\widehat{\Psi}_{4A}^{0.8}$ and $\widehat{\Psi}_{3A}^{0.8}$. Note that the estimators $\widehat{\Psi}_3^{0.5}$, $\widehat{\Psi}_4^{0.5}$, $\widehat{\Psi}_{4A}^{0.5}$, $\widehat{\Psi}_{3A}^{0.5}$, $\widehat{\Psi}_{4A}^{0.8}$ and $\widehat{\Psi}_{3A}^{0.8}$ were introduced in Section 3. Superscripts indicate the value of δ used: $\delta = 0.5$ and $\delta = 0.8$. In what follows, we shall denote the maximal leverage by $h_{\max}$, that is, $h_{\max} = \max(h_1, \ldots, h_n)$. In each simulation scenario, we report the ratio between $h_{\max}$ and $3p/n$; the latter is generally used as a threshold for identifying leverage points. In all simulations, the errors are uncorrelated and normally distributed.

The first numerical evaluation uses the following regression model: $y_i = \beta_1 + \beta_2 x_i + \sigma_i \varepsilon_i$, $i = 1, \ldots, n$. Each random error ε_i has mean zero and unit variance; here, $\sigma_i^2 = \exp\{\alpha x_i\}$. The covariate values are n equally spaced points between zero and one. The sample size is set at $n = 40$. We gradually increase the last covariate value (x_{40}) in order to increase the maximal leverage. We set α at different values in order to vary the heteroskedasticity strength, which we measure as $\lambda = \max\{\sigma_i^2\}/\min\{\sigma_i^2\}$, $i = 1, \ldots, n$. Under homoskedasticity, $\lambda = 1$; otherwise, $\lambda > 1$. The interest lies in testing $\mathcal{H}_0 : \beta_2 = 0$. The null hypothesis is rejected at nominal level η if $\tau = (\widehat{\beta}_2 - 0)^2/\widehat{\text{var}}(\widehat{\beta}_2)$ is greater than $\chi^2_{1,1-\eta}$, the $1 - \eta$ χ^2_1 quantile. The variance estimate in the test statistic denominator is obtained from a heteroskedasticity-consistent covariance matrix estimator.[1] Data generation was performed using $\beta_1 = 1$, $\beta_2 = 0$ and normal errors. All simulations were carried out using the Ox matrix programming language (Doornik, 2006) and are based on 10,000 replications. We use the threshold value $3p/n$ to identify leverage points. The null rejection rates of the different quasi-t tests at the 5% nominal level are presented in Table 1.

The figures in Table 1 show that HC0-based tests can be quite liberal, especially under heavily leveraged data. For instance, under the strongest leverage and $\lambda \approx 49$, the null rejection rate of the test (at the 5% nominal level) exceeds 42%. The same holds true for the test that employs the standard error proposed by Furno (1996), which also incorrectly rejects the null hypothesis in excess of 42%. The HC3 test is liberal and the HC4 test is conservative under leveraged data and heteroskedasticity (9.75% and 3.47%, respectively, when $\lambda \approx 49$ and the leverage ratio equals 3.73). It is noteworthy that the test based on our HC3 variant ($\widehat{\Psi}_3^{0.5}$) outperforms the HC3 test; under the same conditions, its null rejection rate is 5.48%. We also note the excellent finite sample behavior of the test based on another estimator proposed in this paper, namely: $\widehat{\Psi}_{4A}^{0.8}$ (corresponding null rejection rate: 4.18%).

The following regression model was used in our second numerical evaluation: $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{3i}^2 + \sigma_i \varepsilon_i$, $i = 1, \ldots, n$. The ith error has mean zero and unit variance, and $\sigma_i^2 = \exp\{\alpha x_{2i}\}$. The covariate values are obtained as $n = 40$ random draws from the standard lognormal distribution. Data generation was carried out under both homoskedasticity ($\lambda = 1$) and heteroskedasticity ($\lambda \approx 9$ and $\lambda \approx 50$); the errors are normally distributed. Here, $h_{\max}/(3p/n) = 2.475$ (leveraged data). The interest lies in testing $\mathcal{H}_0 : \beta_4 = 0$ against $\mathcal{H}_1 : \beta_4 \neq 0$. The parameter values were set at $\beta_1 = \beta_2 = \beta_3 = 1$ and $\beta_4 = 0$. The null rejection rates of the different quasi-t tests are given in Table 2.

The figures in Table 2 show that the HC0 test is liberal, although not as much as in the previous numerical evaluation. Its null rejection rates at the 5% nominal level range from 9.52% to 11.90%. The test based on Furno's (1996) estimator, $\mathrm{HC0}_R$, displays good finite sample behavior (null rejection rates ranging from 5.53% to 6.66%). It is interesting to note that the HC3 and HC4 tests are quite conservative, especially the latter (null rejection rates around 1%). It is also noteworthy that the best overall performance once again belongs to the test based on $\widehat{\Psi}_{4A}^{0.8}$, whose null rejection rates at the 5% nominal level range from 3.94% to 4.46%. The test based on $\widehat{\Psi}_{4A}^{0.5}$ also performs well.

The next set of simulations is based on the following regression model:

$$y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + \sigma_i \varepsilon_i, \qquad i = 1, \ldots, 50.$$

[1] We also consider the test in which this variance is obtained from the standard OLS estimator, namely $\widehat{\sigma}^2 (X'X)^{-1}$. Note that this test does not have the correct size asymptotically.

Table 1 *Null rejection rates at the 5% nominal level of quasi-t tests with covariate values chosen as a sequence of equally spaced points in the standard unit interval; the last point is gradually increased in order to increase the maximal leverage; here $3p/n = 0.15$*

λ	$h_{\max}/(3p/n)$	OLS	HC0	HC0$_R$	HC3	$\widehat{\Psi}_3^{0.5}$	HC4	$\widehat{\Psi}_4^{0.5}$	$\widehat{\Psi}_{4A}$	$\widehat{\Psi}_{3A}$	$\widehat{\Psi}_{4A}^{0.5}$	$\widehat{\Psi}_{3A}^{0.5}$	$\widehat{\Psi}_{4A}^{0.8}$	$\widehat{\Psi}_{3A}^{0.8}$
1	0.64	4.91	7.2	6.75	5.59	4.68	5.96	5.11	6.41	6.42	5.60	5.60	5.08	5.07
	1.70	4.73	8.09	5.25	5.71	4.44	5.04	3.95	7.66	7.38	6.30	6.04	5.61	5.25
	3.73	4.79	13.75	1.91	7.03	3.24	3.31	1.58	7.87	14.01	5.72	9.80	4.79	8.23
≈ 9	0.64	6.88	7.90	7.84	5.90	4.97	6.37	5.46	6.80	6.80	5.89	5.89	5.33	5.33
	1.70	11.52	10.89	11.29	7.16	5.81	5.70	4.58	9.10	8.99	7.60	7.44	6.80	6.62
	3.73	28.15	30.11	18.00	11.58	6.43	4.81	2.65	9.89	14.10	6.76	9.98	5.28	8.13
≈ 49	0.64	9.89	8.54	8.63	6.48	5.54	6.88	5.98	7.24	7.24	6.40	6.40	5.87	5.87
	1.70	22.36	12.05	19.53	7.06	5.76	4.99	3.90	8.91	8.89	7.29	7.20	6.50	6.40
	3.73	53.62	42.54	42.02	9.75	5.48	3.47	2.21	8.34	11.38	5.31	7.58	4.18	6.10

Table 2 *Null rejection rates at the 5% nominal level of quasi-t tests with covariate values obtained as $n = 40$ random draws from the standard lognormal distribution*; *here* $h_{\max}/(3p/n) = 2.475$

λ	OLS	HC0	HC0$_R$	HC3	$\widehat{\Psi}_3^{0.5}$	HC4	$\widehat{\Psi}_4^{0.5}$	$\widehat{\Psi}_{4A}$	$\widehat{\Psi}_{3A}$	$\widehat{\Psi}_{4A}^{0.5}$	$\widehat{\Psi}_{3A}^{0.5}$	$\widehat{\Psi}_{4A}^{0.8}$	$\widehat{\Psi}_{3A}^{0.8}$
1	5.00	11.90	6.66	3.87	1.74	1.20	0.52	8.70	8.40	5.24	4.94	4.46	4.01
≈ 9	2.91	11.07	6.35	3.73	1.59	1.11	0.48	8.19	8.06	5.16	4.88	4.49	4.05
≈ 50	1.39	9.52	5.53	2.88	1.17	0.93	0.31	7.07	7.00	4.50	4.22	3.94	3.42

Table 3 *Null rejection rates at the 5% nominal level of quasi-t tests with covariate values chosen as per capita spending on public schools in the USA*

n	λ	OLS	HC0	HC0$_R$	HC3	HC4	$\widehat{\Psi}_{4A}^{0.5}$	$\widehat{\Psi}_{4A}^{0.8}$
50	1	4.70	14.01	2.20	5.88	2.02	9.06	5.98
	≈ 9	17.51	25.55	9.13	10.60	3.31	7.55	5.25
	≈ 50	35.64	35.83	19.72	13.45	4.42	7.23	5.09
47	1	4.92	7.58	6.41	5.19	4.86	5.25	4.55
	≈ 9	8.34	9.07	8.05	5.85	5.41	6.13	5.22
	≈ 50	12.42	10.70	10.20	7.08	6.38	7.25	6.53

Each ε_i has mean zero and unit variance; here, $\sigma_i^2 = \exp\{\alpha x_i\}$. The covariate values (values of x_i) are per capita income by state in 1979 in the United States (scaled by 10^{-4}). The response (y) is per capita spending on public schools. The data are presented in Greene (1997, Table 12.1, page 541) and their original source is the U.S. Department of Commerce. Here, however, the response values are generated in the Monte Carlo experiment. That is, we use the observed covariate values, but not the response values. The latter are generated in the numerical exercise. We test $\mathcal{H}_0 : \beta_3 = 0$ (linear specification) against $\mathcal{H}_1 : \beta_3 \neq 0$ (quadratic specification). We consider two situations:

(i) All 50 observations are used; $h_{\max}/(3p/n) = 3.62$.
(ii) Alaska, Washington DC and Mississippi (the three leverage points) are removed from the data ($n = 47$); $h_{\max}/(3p/n) = 1.09$.

In both cases, the random errors are normally distributed. Hereafter, we shall focus on the following tests: OLS, HC0, HC0$_R$, HC3, HC4, $\widehat{\Psi}_{4A}^{0.5}$ and $\widehat{\Psi}_{4A}^{0.8}$. Their null rejection rates are given in Table 3.

We note from the results reported in Table 3 that the HC0 test is quite liberal, especially under strong heteroskedasticity and leveraged data and that Furno's (1996) variant, HC0$_R$, is also liberal. When $n = 50$ and $\lambda \approx 50$, their empirical sizes at the 5% nominal level are, respectively, 35.83% and 19.72%. These are quite large size distortions. These tests are also liberal even when the three leverage points are not in the data and heteroskedasticity is intense ($\lambda \approx 50$); both tests reject the null hypothesis in excess of 10% (i.e., more than twice the nominal size). Under leveraged data, the HC3 test is liberal and the HC4 test is conservative. For instance, the HC3-based test null rejection rate when $n = 50$ and $\lambda \approx 50$ equals 13.45%. Once again, the test with the best overall performance is that based on $\widehat{\Psi}_{4A}^{0.8}$: its null rejection rates range from 5.09% to 5.98% when $n = 50$ (leverage points in the data) and from 4.55% to 6.53% when the $n = 47$ (leverage points removed from the data). The test based on $\widehat{\Psi}_{4A}^{0.5}$ performs well under unequal error variances.

Next, we shall consider nonnormal data generating processes. The model is

$$y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \sigma_i \varepsilon_i,$$

where ε_i has mean zero, variance one and is uncorrelated with ε_j for all $i \neq j$. Additionally, $\sigma_i^2 = \exp\{\alpha x_{2i}\}$. We test $\mathcal{H}_0 : \beta_4 = 0$ against $\mathcal{H}_1 : \beta_4 \neq 0$ at the 5% nominal level. The sample size is set at 40 and the covariate values are obtained as random draws from the standard lognormal distribution. We consider the following distributions for ε_i: t_5 (fat tailed, normalized to have unit variance), χ_5^2 (asymmetric, normalized to have unit variance) and exponential with unit mean (asymmetric, normalized to have zero mean). The null rejection rates of the OLS, HC0, HC0$_R$, HC3, HC4, $\widehat{\Psi}_{4A}^{0.5}$ and $\widehat{\Psi}_{4A}^{0.8}$ tests for $p = 4$ and $p = 6$ are given in Tables 4 and 5, respectively. The figures in these tables show that the tests rejection rates are approximately the same under the three error distributions. They also show that the HC0 and HC0$_R$ tests are slightly more oversized when $p = 6$ relative to the case where $p = 4$. It is also noteworthy that the HC4 test is quite undersized (conservative) in both cases ($p = 4$ and $p = 6$): its null rejection rates never exceeds 2%, more often being around 1%. We also note that the HC3 is undersized. The best performing test is that based on standard errors obtained from $\widehat{\Psi}_{4A}^{0.5}$, the consistent estimator we proposed in Section 3; the $\widehat{\Psi}_{4A}^{0.8}$ is the runner-up. For instance, when the errors are t-distributed, heteroskedasticity is moderate (strong) and $p = 6$, the

Table 4 *Null rejection rates at the 5% nominal level of quasi-t tests with covariate values obtained as $n = 40$ random draws from the standard lognormal distribution; here $p = 4$ and $h_{\max}/(3p/n) = 2.475$*

errors	λ	OLS	HC0	HC0$_R$	HC3	HC4	$\widehat{\Psi}_{4A}^{0.5}$	$\widehat{\Psi}_{4A}^{0.8}$
t_5	1	5.15	11.73	6.61	3.79	1.07	5.45	4.43
	≈ 9	3.01	10.72	6.13	3.62	0.98	5.15	4.16
	≈ 50	1.58	9.15	5.44	2.89	0.85	4.60	3.87
χ_5^2	1	4.93	11.27	6.46	3.55	0.96	4.98	4.02
	≈ 9	2.93	10.38	6.03	3.43	1.03	4.88	4.13
	≈ 50	1.65	8.75	5.27	2.64	0.75	4.04	3.44
exponential	1	5.36	11.01	6.52	3.17	0.88	4.55	3.67
	≈ 9	3.28	10.56	6.22	3.13	0.87	4.70	3.82
	≈ 50	1.82	9.74	5.72	2.93	0.81	4.35	3.89

Table 5 *Null rejection rates at the 5% nominal level of quasi-t tests with covariate values obtained as $n = 40$ random draws from the standard lognormal distribution; here $p = 6$ and* $h_{\max}/(3p/n) = 1.6535$

errors	λ	OLS	HC0	HC0$_R$	HC3	HC4	$\widehat{\Psi}_{4A}^{0.5}$	$\widehat{\Psi}_{4A}^{0.8}$
t_5	1	5.27	11.98	8.60	3.92	1.52	5.02	3.91
	≈ 9	3.23	11.15	8.20	3.47	1.34	4.85	3.86
	≈ 50	1.66	9.72	6.72	2.75	1.10	3.94	3.19
χ_5^2	1	5.27	12.27	8.66	3.54	0.97	4.88	3.71
	≈ 9	3.46	10.88	7.79	3.36	0.96	4.37	3.36
	≈ 50	1.89	8.86	6.19	2.60	0.85	3.44	2.62
exponential	1	5.40	11.05	7.93	3.14	0.86	4.26	3.07
	≈ 9	3.58	10.44	7.38	3.06	0.86	4.01	3.04
	≈ 50	1.96	9.60	6.57	2.64	0.73	3.72	2.80

null rejection rate of the test based on $\widehat{\Psi}_{4A}^{0.5}$ is 4.85% (3.94%) whereas those of the HC3 and HC4 tests are 3.47% and 1.34% (2.75% and 1.10%). The test based on $\widehat{\Psi}_{4A}^{0.8}$ also performs well.

5 Concluding remarks

This paper addressed the issue of performing testing inference in linear regressions under heteroskedasticity of unknown form. The most commonly used heteroskedasticity-robust test is that based on standard errors obtained from White's (1980) estimator, HC0. Such a test tends to be quite liberal in small samples, especially when the data contain points of high leverage. Tests based on the HC3 and HC4 estimators are generally viewed as good alternatives. However, our numerical results show that they may not deliver accurate inferences in small samples; tests based on the former tend to be liberal whereas those based on the latter tend to be conservative. Our numerical results have also shown that testing inference based on Furno's (1996) estimator may be misleading since her test can be quite liberal in some situations. In this paper, we proposed two new classes of heteroskedasticity-robust standard errors. They are based on the heteroskedasticity-consistent covariance matrix estimators of Cribari-Neto and Lima (2010). The numerical evidence in Section 4 favored two of the estimators we proposed, namely: $\widehat{\Psi}_{4A}^{0.5}$ and $\widehat{\Psi}_{4A}^{0.8}$. The corresponding tests displayed superior small sample performance in the Monte Carlo simulations. We strongly recommend that practitioners base their inferences on such tests.

Acknowledgments

The authors thank an anonymous referee for comments and suggestions. Francisco Cribari-Neto gratefully acknowledges partial financial support from CNPq.

References

Chesher, A. and Jewitt, I. (1987). The bias of a heteroskedasticity consistent covariance matrix estimator. *Econometrica* **55**, 1217–1222. MR0913369

Cribari-Neto, F. (2004). Asymptotic inference under heteroskedasticity of unknown form. *Computational Statistics and Data Analysis* **45**, 215–233. MR2045469

Cribari-Neto, F., Ferrari, S. L. P. and Cordeiro, G. M. (2000). Improved heteroscedasticity-consistent covariance matrix estimators. *Biometrika* **87**, 907–918. MR1813983

Cribari-Neto, F. and Lima, M. G. A. (2010). Sequences of bias-adjusted covariance matrix estimators under heteroskedasticity of unknown form. *Annals of the Institute of Statistical Mathematics* **62**, 1053–1082. MR2729154

Cribari-Neto, F., Souza, T. C. and Vasconcellos, K. L. P. (2007). Inference under heteroskedasticity and leveraged data. *Communications in Statistics, Theory and Methods* **36**, 1877–1888. Errata: **37** (2008) 3329–3330. MR2409817

Davidson, R. and MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford: Oxford Univ. Press. MR1350531

Doornik, J. A. (2006). *An Object-oriented Matrix Programming Language—Ox 4*. London: Timberlake Consultants Press. Available at http://www.doornik.com.

Furno, M. (1996). Small sample behavior of a robust heteroskedasticity consistent covariance matrix estimator. *Journal of Statistical Computation and Simulation* **54**, 115–128.

Greene, W. H. (1997). *Econometric Analysis*, 3rd ed. Upper Saddle River: Prentice Hall.

Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics* **19**, 285–292. MR0458734

Horn, S. D., Horn, R. A. and Duncan, D. B. (1975). Estimating heteroskedastic variances in linear models. *Journal of the American Statistical Association* **70**, 380–385. MR0370946

Long, J. S. and Ervin, L. H. (2000). Using heteroskedasticity-consistent standard errors in the linear regression model. *The American Statistician* **54**, 217–224.

Qian, L. and Wang, S. (2001). Bias-corrected heteroscedasticity robust covariance matrix (sandwich) estimators. *Journal of Statistical Computation and Simulation* **70**, 161–174. MR1869718

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* **48**, 817–838. MR0575027

Departamento de Estatística
Universidade Federal de Pernambuco
Cidade Universitária
Recife/PE, 50740–540
Brazil
E-mail: cribari@de.ufpe.br
gloria.abage@gmail.com
URL: http://www.de.ufpe.br/~cribari