ON THE CORRECTIONS TO INFORMATION MATRIX TESTS

Francisco Cribari–Neto
Department of Economics, Southern Illinois University
Carbondale IL, 62901-4515, USA

(Forthcoming in *Econometric Reviews*)

ABSTRACT

This paper addresses the issue of designing finite-sample corrections to information matrix tests. We review a Cornish-Fisher correction that has been proposed elsewhere and propose an alternative, Bartlett-type correction. Simulation results for skewness, excess kurtosis, normality and heteroskedasticity tests are given.

Keywords and Phrases: Bartlett correction; Cornish-Fisher expansion; Edgeworth expansion; heteroskedasticity test; information matrix test; normality test; size-correction.

Journal of Economic Literature Classification: C52, C12.

1. Introduction

The information matrix (IM) test was introduced by White (1982) as a general test for model specification. Many important tests can also be obtained as special cases of the class of IM tests, such as: skewness, excess kurtosis, normality and heteroskedasticity tests. However, the available Monte Carlo simulation results on the IM test have shown that its size performance can be quite distorted in finite samples, especially when certain asymptotic variance estimators are used (see Chesher and Spady, 1991; Davidson and MacKinnon, 1992; Orme, 1990; Taylor, 1987; among others). As stated by Taylor (1987, p.66): "The Monte Carlo results then indicate that some type of finite sample alterations could be very valuable for the IM test. Corrections are indeed necessary before any formal evaluations of the power properties of the test can be undertaken." This is the motivation for this paper.

It has been shown by Chesher (1984) that the IM test has a Lagrange multiplier interpretation, and this allowed Chesher and Spady (1991) to use Harris' (1985, 1987) results to give an asymptotic expansion for the null distribution of the IM test statistic to order n^{-1}, where n is the sample size. They have also shown how to obtain transformed critical values to be used in the test to improve its finite-sample behavior. In this paper we argue that Chesher and Spady's results can also be used to obtain a transformed IM test statistic. The 'improved' test can be carried out by comparing the transformed statistic with critical values obtained from chi-squared reference tables.

The plan of the paper is as follows. Section 2 briefly reviews the information matrix test. Chesher and Spady's size correction is described in Section 3. This section also proposes a new correction based on a transformation of the original statistic. Sections 4 and 5 give simulation results for skewness, excess kurtosis, normality and heteroskedasticity tests, and section 6 examines some further issues. Concluding remarks are given in the last section.

2. The information matrix test

Let $y_1, \ldots, y_n$ be a set of independent vectors of observations with true density $g(y)$. Suppose we fit a model to the data using a family of parametric density functions $f(y;\theta)$, where θ is a p-vector of parameters, and assume that $g(y)$ and $f(y;\theta)$ satisfy the regularity conditions stated by White (1982). Let $\hat{\theta}$ be the quasi-maximum likelihood estimator and θ_g the value of θ that maximizes the Fraser (1965) information of $f(y;\theta)$ under $g(y)$ which is defined as $\int \log[f(y;\theta)]g(y)dy$. It was shown by White (1982) that under some regularity conditions $\hat{\theta}$ exists and $\hat{\theta} \xrightarrow{a.s.} \theta_g$. If the model is correctly specified, $\hat{\theta}$ is a consistent estimator of the true parameter vector. Moreover, in the absence of misspecification (*i.e.*, $f(\cdot;\theta_0) = g(\cdot)$ for some θ_0), the information matrix equality holds, that is,

$$J(\theta) := \mathrm{E}\left[\frac{\partial \log f(y;\theta)}{\partial\theta}\frac{\partial \log f(y;\theta)}{\partial\theta'}\right] = -\mathrm{E}\left[\frac{\partial^2 \log f(y;\theta)}{\partial\theta\,\partial\theta'}\right] =: K(\theta),$$

for $\theta = \theta_0$. In the case of misspecification, however, this equality need not hold when expectations are taken with respect to $g(y)$. This is the basis for the class of tests introduced by White (1982) and known as information matrix tests.

White's test can be described as follows. Define

$$\varphi(y;\theta) := \mathrm{vech}\left(\frac{\partial^2 \log f(y;\theta)}{\partial\theta\,\partial\theta'} + \frac{\partial \log f(y;\theta)}{\partial\theta}\frac{\partial \log f(y;\theta)}{\partial\theta'}\right),$$

$$\bar{\varphi}(\theta) := n^{-1}\sum_{i=1}^{n}\varphi(y_i;\theta),$$

$$\Psi(\theta) := n^{-1}\sum_{i=1}^{n}\frac{\partial\varphi(y_i;\theta)}{\partial\theta'},$$

$$s(y_i;\theta) := \frac{\partial \log f(y_i;\theta)}{\partial\theta}, \quad i = 1, \ldots, n.$$

Then, the IM test statistic is given by

$$IM = n\,\bar{\varphi}(\hat{\theta})'\,\Lambda^{-1}(\hat{\theta})\,\bar{\varphi}(\hat{\theta}), \tag{1}$$

where

$$\Lambda(\hat{\theta}) := n^{-1}\sum_{i=1}^{n}[\varphi(y_i;\hat{\theta}) - \Psi(\hat{\theta})K^{-1}(\hat{\theta})s(y_i;\hat{\theta})][\varphi(y_i;\hat{\theta}) - \Psi(\hat{\theta})K^{-1}(\hat{\theta})s(y_i;\hat{\theta})]'.$$

For the remainder of the paper, it will be assumed that the expected values in the expression above are taken with $\hat{\theta}$ substituted for θ; see Chesher and Spady (1991, p.789).

Under the null hypothesis and the regularity conditions stated by White (1982), IM is asymptotically distributed as χ^2_s, where s is the number of restrictions under test. The full information matrix test is carried out by taking $s = p\,(p+1)/2$. It is noteworthy that several important tests can be obtained as special cases of this class of tests, as for instance the Bera-Jarque normality test (Jarque and Bera, 1980, 1987), tests for skewness and excess kurtosis, and a version of White's (1980) heteroskedasticity test.

3. Finite–sample corrections to information matrix tests

As shown by Chesher (1984), the information matrix test has a Lagrange multiplier interpretation when one allows the p-vector of parameters θ to be non-fixed and tests the null hypothesis that θ is not random. This interpretation allowed Chesher and Spady (1991) to use Harris' (1985, 1987) asymptotic expansion for the null distribution function of the IM test statistic defined in (1) since this statistic is a special case of the class of Lagrange multiplier tests. Harris' expansion is given by

$$\begin{aligned}\Pr[IM \le u] = H_s(u) + (24\,n)^{-1}\{\alpha_3 H_{s+6}(u) + (\alpha_2 - 3\alpha_3)H_{s+4} + (3\alpha_3 - 2\alpha_2 \\ + \alpha_1)H_{s+2}(u) + (\alpha_2 - \alpha_1 - \alpha_3)H_s\} + o(n^{-1}),\end{aligned} \qquad (2)$$

where $H_m(\cdot)$ denotes the distribution function of a χ^2_m random variable. The α's are functions of the joint cumulants of log-likelihood derivatives. For a definition of these α's, see Harris (1985, 1987); for matrix expressions, see Ferrari and Cordeiro (1994).

Harris (1985) also provided a Cornish-Fisher expansion using Hill and Davis' (1968) inverse formula. That is, he provides a solution to the inversion of the Edgeworth expansion in terms of the quantiles of the second order distribution. The modified quantiles can be written as

$$\begin{aligned}\tilde{u}_\alpha = u_\alpha + (12\,n)^{-1}\Bigg\{ & \frac{\alpha_3 u_\alpha[u_\alpha^2 + (s+4)u_\alpha + (s+2)(s+4)]}{s(s+2)(s+4)} \\ & \frac{u_\alpha(u_\alpha + s + 2)(\alpha_2 - 3\alpha_3)}{s(s+2)} + \frac{u_\alpha(3\alpha_3 - 2\alpha_2 + \alpha_1)}{s}\Bigg\} + o(n^{-1}),\end{aligned} \qquad (3)$$

where u_α is the $(1-\alpha)$th quantile of a chi-squared distribution with s degrees of freedom, *i.e.*, $\Pr[\chi^2_s \ge u_\alpha] = \alpha$.

It is possible to write the expansion in (2) in a more convenient way. To this end, the following recurrence relations will be useful (Patel, Kapadia and Owen, 1976, p.221):

$$\begin{aligned}H_{r+2}(u) &= H_r(u) - 2ur^{-1}h_r(u), \\ h_{r+2}(u) &= ur^{-1}h_r(u),\end{aligned}$$

where $h_r(u)$ denotes the density function of a χ^2_r variate and r is a positive integer. From the relations above we obtain:

$$\begin{aligned} h_{r+4}(u) &= \frac{u^2}{r(r+2)} h_r(u), \\ h_{r+6}(u) &= \frac{u^3}{r(r+2)(r+4)} h_r(u), \\ H_{r+4}(u) &= H_r(u) - 2\left\{ \frac{u}{r} + \frac{u^2}{r(r+2)} \right\} h_r(u), \\ H_{r+6}(u) &= H_r(u) - 2\left\{ \frac{u}{r} + \frac{u^2}{r(r+2)} + \frac{u^3}{r(r+2)(r+4)} \right\} h_r(u). \end{aligned}$$

Making use of these relations, we can rewrite (2), after some algebra, as

$$\Pr[IM \leq u] = H_s(u) - h_s(u) \frac{1}{n} \sum_{j=1}^{3} \gamma_j \, u^j + o(n^{-1}), \tag{4}$$

where

$$\gamma_1 := \frac{\alpha_1 - \alpha_2 + \alpha_3}{12s}, \tag{5a}$$

$$\gamma_2 := \frac{\alpha_2 - 2\alpha_3}{12s(s+2)}, \tag{5b}$$

$$\gamma_3 := \frac{\alpha_3}{12s(s+2)(s+4)}. \tag{5c}$$

It is more convenient to write the expansion for the null distribution function of the IM statistic as in (4) than as given by Chesher and Spady (1991)—equation (2)—, since the expression above does not involve distribution functions of chi-squared variates with different degrees of freedom.

It is interesting to notice that by making use of the notation introduced above we can rewrite the Cornish-Fisher expansion in (3) as

$$\tilde{u}_\alpha = u_\alpha \left[1 + \frac{1}{n} \sum_{j=1}^{3} \gamma_j \, u_\alpha^{j-1} \right] + o(n^{-1}). \tag{6}$$

Equation (6) makes clear that one can obtain improved critical values (to order n^{-1}) for information matrix tests by multiplying the tabulated critical values by the bracketed factor.

We shall now show how to design improved IM tests based on transformed test statistics rather than on transformed critical values. To this end, we shall follow Chesher and Spady (1991) and recall that the information matrix test can be interpreted as a Lagrange multiplier test where the null hypothesis states that the p-vector of parameters θ (or a subset of the elements in this vector) is not subject to random variation.

Given the interpretation of the information matrix test as a Lagrange multiplier test, it can be shown that the modified statistic

$$\widetilde{IM} = IM\left[1 - \frac{1}{n}\sum_{j=1}^{3} \zeta_j \, IM^{j-1}\right] \tag{7}$$

has a chi-squared null distribution to order n^{-1} if and only if $\zeta_j = \gamma_j$, $j = 1, 2, 3$, where the γ's are as before, *i.e.*, as defined in (5a)–(5c). The proof follows from the results in Cordeiro and Ferrari (1991). That is, the modified statistic $\widetilde{IM}$ is distributed as chi-squared to order n^{-1}: $\Pr[\widetilde{IM} \leq u] = \Pr[\chi_s^2 \leq u] + o(n^{-1})$, that is, $\Pr[\widetilde{IM} \leq u] = \Pr[\chi_s^2 \leq u]$ when terms of order smaller than n^{-1} are ignored. The correction based on the modified statistic in (7) is known as the 'Bartlett-type correction'.

It is interesting to note the similarity between (6) and (7). In (6) a correction is obtained by multiplying the tabulated critical value by [1 + polynomial on u], while in (7) a correction to the test statistic involves multiplying the unmodified statistic by [1 − polynomial on IM]. It is noteworthy that the two corrected tests, *i.e.*, the one based on the corrected critical value and the one based on the corrected statistic proposed here, are equivalent to order n^{-1} under the null hypothesis: $\Pr[IM \leq \tilde{u}] = \Pr[\widetilde{IM} \leq u]$ when terms of order smaller than n^{-1} are neglected.

It can be also shown that α_3 is always nonnegative by assuming, without loss of generality, that Fisher's total information can be written as $n\,I$, where I is the identity matrix (see Cordeiro, 1987, p.268), which implies that α_3 can be written as a sum of nonnegative terms. Since $\alpha_3 \geq 0$, large values of IM can lead to negative values of $\widetilde{IM}$, and this may seem an undesirable feature at first glance. However, this will only happen for very large values of IM and for such values there is already strong evidence against the null hypothesis and hence no correction is needed. Consider first the case where $\alpha_3 = 0$. If $\alpha_2 > 0$, $\widetilde{IM}$ will be nonnegative whenever $IM \leq b$, where $b = [12\,n\,s(s+2)/\alpha_2] + (s+2)(1 - \alpha_1/\alpha_2) \approx 12\,n\,s(s+2)/\alpha_2$; the approximation follows from neglecting the terms of order $O(1)$ which are dominated by the term that grows with n for sufficiently large n. Since IM is asymptotically distributed as chi-squared, for the typical sample sizes used in econometrics $IM > b$ will happen only when there is strong evidence against the null hypothesis, and in this case the correction is not necessary. If $\alpha_2 < 0$, $\widetilde{IM}$ will be nonnegative whenever $IM \geq b$. Since IM is itself nonnegative, this condition will always be fulfilled. When $\alpha_3 > 0$, it can be shown using similar arguments that $\widetilde{IM}$ will be negative only for large values of IM and in this case, as noted before, no correction is needed.

The discussion above leads us to propose a hybrid statistic:

$$\widehat{IM} = \begin{cases} \widetilde{IM} & \text{if } IM \leq \nu_s \\ IM & \text{otherwise,} \end{cases}$$

where ν_s can be chosen as, say, the 99.5% upper quantile of the χ_s^2 distribution. Once again, the rationale here is that for large values of the test statistic one has strong evidence against the hypothesis under test and hence there is no need for correcting the test.

Another solution is to use an alternative correction which is equivalent to (7) to order n^{-1} under the null hypothesis and always delivers nonnegative test statistics. A natural candidate is

$$IM^{+} = IM \exp\left\{ -\frac{1}{n}\sum_{j=1}^{3} \zeta_j \, IM^{j-1} \right\}.$$

Both Bartlett-corrected statistics will be considered in the Monte Carlo simulations reported in the remaining sections.

It is interesting to note that it is much harder to propose a hybrid version—similar to the one above—for the test based on the size-corrected critical value given in (6) in order to avoid the possibility of obtaining negative transformed critical values. This would be of the form

$$\hat{u} = \begin{cases} \tilde{u} & \text{if } |u - \tilde{u}|/u \le \nu_c \\ u & \text{otherwise.} \end{cases}$$

However, there are some difficulties involved in doing this. First, a large value of $|u - \tilde{u}|/u$ is not evidence that a correction is not necessary. Second, there is no obvious criterion for choosing ν_c. Finally, the truncation does not depend upon the evidence against H_0 conveyed by the sample.

4. Normality tests

We shall now present some Monte Carlo simulation results on the performance of skewness, excess kurtosis and normality tests, as examples of IM-type tests. The normality test is commonly referred to as the Bera-Jarque test (Jarque and Bera, 1980, 1987). The tests are applied to the residuals from the regression

$$y_i = 2.5 + 5\,x_i + \varepsilon_i, \quad i = 1, \ldots, n,$$

where x is chosen as the quantiles of a U(10, 20) distribution, and $\varepsilon_i \sim \text{iid}(0, \sigma^2)$. The null hypothesis under test is $H_0 : \varepsilon \sim N$. Pseudo-random numbers were generated from a standard normal distribution. All simulations are based on 10,000 replications. The rejection rates for the original tests and their three corrected versions under the null hypothesis are given in Tables 1, 2 and 3 for the skewness, excess kurtosis and normality tests, respectively. Entries are percentages.

The figures in Tables 1, 2 and 3 confirm the well known tendency of the three tests to reject the null hypothesis under normality less often than expected based on the nominal size of the test. The results in Table 1 (skewness) show that all corrections lead to substantial improvements in the size behavior of the test. The hybrid corrected statistic $\widehat{IM}$ outperformed the other corrections when the sample size was small. Overall, the corrections seem to work properly. An interesting result conveyed in Tables 2 and 3 (excess kurtosis and normality) is that none of the corrected tests seemed to deliver a substantial improvement in the finite-sample behavior of the uncorrected test. Some improvement occurs for $n \ge 100$. For smaller

Table 1: Skewness Test

n	α	(1)	(2)	(3)	(4)	(5)
10	10	3.2	12.3	22.8	20.5	0.6883
	5	1.3	5.7	12.7	3.8	0.3476
	1	0.1	0.0	0.0	0.0	0.4905
20	10	5.7	10.7	12.6	11.3	0.3442
	5	2.9	4.9	5.5	4.3	0.1738
	1	0.6	0.3	0.0	0.3	0.2452
30	10	6.7	10.7	11.4	10.8	0.2295
	5	3.4	4.8	4.9	4.3	0.1159
	1	0.7	0.4	0.0	0.4	0.1635
40	10	7.4	10.1	10.4	10.1	0.1721
	5	3.6	4.6	4.7	4.4	0.0869
	1	0.7	0.4	0.2	0.5	0.1226
50	10	8.4	10.7	10.9	10.7	0.1377
	5	4.3	5.1	5.1	4.9	0.0695
	1	1.0	0.5	0.6	0.8	0.0981
60	10	8.3	10.3	10.4	10.3	0.1147
	5	4.1	4.7	4.8	4.6	0.0579
	1	1.0	0.7	0.7	0.8	0.0817
80	10	8.6	10.0	10.0	10.0	0.0860
	5	4.3	4.8	4.8	4.7	0.0435
	1	1.0	0.8	0.8	0.8	0.0613
100	10	9.0	10.2	10.2	10.2	0.0688
	5	4.8	5.1	5.1	5.1	0.0348
	1	1.2	1.0	1.0	1.0	0.0491
150	10	9.5	10.2	10.2	10.2	0.0459
	5	4.6	4.8	4.8	4.8	0.0232
	1	0.9	0.7	0.7	0.7	0.0327
200	10	9.5	10.1	10.2	10.1	0.0344
	5	5.3	5.5	5.5	5.5	0.0174
	1	1.2	1.1	1.1	1.1	0.0245
250	10	9.1	9.5	9.5	9.5	0.0275
	5	4.7	4.9	4.9	4.9	0.0139
	1	1.0	0.9	0.9	0.9	0.0196

Note: (1) denotes $\Pr[IM_N \geq u]$, (2) denotes $\Pr[\widehat{IM_N} \geq u]$, (3) denotes $\Pr[IM_N^+ \geq u]$, (4) denotes $\Pr[IM_N \geq \tilde{u}]$ and (5) denotes $|u - \tilde{u}|/u$; here, $\nu_s = 8$.

Table 2: Excess Kurtosis Test

n	α	(1)	(2)	(3)	(4)	(5)
10	10	0.3	32.0	82.4	100.0	5.3397
	5	0.1	19.2	79.1	100.0	3.2630
	1	0.0	3.6	72.9	0.0	8.4298
20	10	2.0	21.7	52.1	100.0	2.6698
	5	1.3	12.1	44.8	100.0	1.6315
	1	0.5	3.1	32.1	0.0	4.2149
30	10	2.7	18.1	37.2	100.0	1.7799
	5	1.8	9.6	27.8	100.0	1.0877
	1	0.8	2.5	14.9	0.0	2.8099
40	10	3.1	15.4	27.3	100.0	1.3349
	5	2.0	8.3	18.5	29.2	0.8158
	1	0.9	1.6	7.5	0.1	2.1075
50	10	4.0	15.0	22.9	100.0	1.0679
	5	2.5	7.8	14.7	14.3	0.6526
	1	1.3	1.0	4.3	0.2	1.6860
60	10	4.4	14.4	20.1	54.8	0.8899
	5	2.7	7.5	12.1	9.8	0.5438
	1	1.2	0.9	1.7	0.3	1.4050
80	10	5.2	12.8	16.1	27.6	0.6675
	5	3.2	6.5	8.4	7.2	0.4079
	1	1.4	1.1	0.0	0.3	1.0537
100	10	5.4	11.9	13.8	19.2	0.5340
	5	2.9	5.7	6.6	5.8	0.3263
	1	1.3	0.9	0.0	0.4	0.8430
150	10	6.4	11.1	11.7	13.7	0.3560
	5	3.3	5.2	5.1	5.3	0.2175
	1	1.4	1.0	0.0	0.6	0.5620
200	10	7.3	11.3	11.4	12.6	0.2670
	5	3.8	5.3	4.9	5.3	0.1632
	1	1.3	1.0	0.0	0.7	0.4215
250	10	7.1	10.6	10.5	11.3	0.2136
	5	3.6	4.9	4.9	4.8	0.1305
	1	1.3	0.9	0.9	0.7	0.3372

Note: (1) denotes $\Pr[IM_N \geq u]$, (2) denotes $\Pr[\widehat{IM_N} \geq u]$, (3) denotes $\Pr[IM_N^+ \geq u]$, (4) denotes $\Pr[IM_N \geq \tilde{u}]$ and (5) denotes $|u - \tilde{u}|/u$; here, $\nu_s = 8$.

Table 3: The Bera-Jarque Normality Test

n	α	(1)	(2)	(3)	(4)	(5)
10	10	1.6	26.1	83.1	100.0	3.2738
	5	0.8	15.5	79.1	2.9	0.4346
	1	0.2	8.0	69.9	0.0	11.8557
20	10	3.8	20.1	50.1	100.0	1.6369
	5	2.6	12.8	41.8	3.7	0.2173
	1	1.1	5.7	27.3	0.0	5.9279
30	10	4.6	17.9	35.1	100.0	1.0913
	5	3.0	11.1	26.6	3.9	0.1449
	1	1.4	3.0	13.5	0.0	3.9519
40	10	4.9	16.2	26.7	58.1	0.8185
	5	3.2	9.5	18.6	3.9	0.1086
	1	1.5	1.1	6.7	0.1	2.9639
50	10	5.7	15.4	21.9	31.8	0.6548
	5	3.8	8.5	14.1	4.3	0.0869
	1	1.8	1.4	0.0	0.1	2.3711
60	10	5.8	14.1	19.1	22.5	0.5456
	5	3.7	7.4	10.8	4.3	0.0724
	1	1.8	1.4	0.0	0.2	1.9760
80	10	6.4	12.8	15.1	16.0	0.4092
	5	4.0	6.0	6.8	4.4	0.0543
	1	1.9	1.6	0.0	0.2	1.4820
100	10	6.6	11.9	12.9	13.7	0.3274
	5	4.1	4.7	4.7	4.5	0.0435
	1	1.9	1.4	0.0	0.3	1.1856
150	10	7.2	11.0	11.2	11.7	0.2183
	5	4.3	3.6	2.6	4.5	0.0290
	1	1.7	1.3	0.0	0.4	0.7904
200	10	8.2	11.1	10.9	11.5	0.1637
	5	4.6	3.7	2.7	4.8	0.0217
	1	1.8	1.4	0.0	0.5	0.5928
250	10	7.8	9.9	9.5	10.0	0.1310
	5	4.2	3.4	2.7	4.3	0.0174
	1	1.6	1.1	0.0	0.6	0.4742

Note: (1) denotes $\Pr[IM_N \geq u]$, (2) denotes $\Pr[\widehat{IM_N} \geq u]$, (3) denotes $\Pr[IM_N^+ \geq u]$, (4) denotes $\Pr[IM_N \geq \tilde{u}]$ and (5) denotes $|u - \tilde{u}|/u$; here, $\nu_s = 10.5$.

sample sizes the corrected tests display very poor behavior. In particular, the test based on transformed critical values yielded estimated sizes of 100% for the nominal levels of 10% and 5% in some cases. For $n \geq 100$ the correction to the Bera-Jarque test (Table 3) based on the the hybrid transformation of the IM statistic works better for the nominal size of 10%, while the correction based on the transformation of critical values is more effective for $\alpha = 5\%$. For the excess kurtosis test (Table 2), the hybrid corrected statistic outperformed the other two corrected tests. For small sample sizes, none of the corrections work well and Chesher and Spady's correction has a dismal effect on the original test. For moderately large sample sizes, the corrections become more effective. Overall, the convergence of the IM statistic to its asymptotic chi-squared distribution is very slow when it comes to excess kurtosis and normality tests.

5. A heteroskedasticity test

We shall now turn to the heteroskedasticity test proposed by Breusch and Pagan (1979) and Godfrey (1978). The auxiliary variables for this test are chosen as in White's (1980) test. Therefore, the version of the test used here is not the Studentized version of White (1980) and Koenker (1981). It is the Breusch-Pagan-Godfrey test with the choice of exogenous auxiliary variables as in White's test. Finite-sample corrections to this test were studied by Honda (1988) and Cribari–Neto and Ferrari (1995). They are evaluated in more detail here. It should be said that the numerical values for α_1, α_2 and α_3 given by Chesher and Spady (1991) are not correct. The values of the α's used here were obtained using the formulas in Honda (1988) and Cribari–Neto and Ferrari (1995). Unlike the previous case (skewness, excess kurtosis and normality), the α's for the heteroskedasticity test are covariate dependent.

The regression model considered here is

$$y_i = 2.5 + 5\,x_i + \varepsilon_i, \;\; i = 1, \ldots, n,$$

where $\varepsilon_i \sim \mathrm{NID}(0, \sigma_i^2)$ with $\sigma_i^2 = h(\delta_0 + \delta_1\, x_i + \delta_2\, x_i^2)$, h being an unknown, twice continuously differentiable function, usually called 'the skedastic function'. The null hypothesis under test is $\mathrm{H}_0 : \delta_1 = \delta_2 = 0$, under which $\sigma_i^2 = h(\delta_0) = \sigma^2$, that is, the error variance is constant across observations. The alternative hypothesis is $\mathrm{H}_1 : \max\{|\delta_1|, |\delta_2|\} \neq 0$. The covariate values are 10 values evenly spaced in $[-1, +1]$ for $n = 10$. These values are replicated for other sample sizes. This covariate choice is similar to a simulation design used by Chesher and Spady (1991). Rejection rates based on 10,000 replications using $h(\delta_0) = 1$ and $\delta_1 = \delta_2 = 0$ are given in Table 4. Entries are percentages.

These simulation results show that the three corrected IM tests can improve the finite-sample performance of the original test. It is noteworthy that the finite-sample corrections were much more effective in the context of testing for heteroskedasticity than in the normality test. Also, the two Bartlett-corrected statistics displayed similar small sample performance.

Table 4: Heteroskedasticity Test

n	α	(1)	(2)	(3)	(4)	(5)
10	10	3.6	7.2	7.7	10.3	0.3166
	5	1.6	3.7	4.1	5.2	0.3232
	1	0.3	0.9	1.0	0.5	0.1037
20	10	6.6	9.3	9.4	10.1	0.1583
	5	3.1	4.8	4.9	5.1	0.1616
	1	0.8	1.1	1.1	1.0	0.0519
30	10	7.6	9.7	9.8	10.1	0.1055
	5	3.6	4.8	4.9	4.9	0.1077
	1	0.8	0.9	0.8	0.8	0.0346
40	10	7.8	9.3	9.4	9.5	0.0792
	5	3.6	4.7	4.8	4.8	0.0808
	1	0.8	0.9	0.9	0.9	0.0259
50	10	8.6	10.0	10.0	10.2	0.0633
	5	3.9	4.9	4.9	5.0	0.0646
	1	0.7	0.7	0.7	0.7	0.0208
60	10	8.6	9.9	9.9	10.0	0.0528
	5	4.2	4.8	4.9	4.9	0.0539
	1	0.9	1.0	1.0	1.0	0.0173
80	10	9.3	10.2	10.2	10.2	0.0396
	5	4.4	5.0	5.0	5.0	0.0404
	1	0.9	1.0	1.0	0.9	0.0130
100	10	9.4	10.2	10.2	10.2	0.0317
	5	4.4	4.9	4.9	4.9	0.0323
	1	1.0	1.0	1.0	1.0	0.0104
150	10	9.8	10.3	10.3	10.3	0.0211
	5	4.9	5.3	5.3	5.3	0.0216
	1	0.8	0.8	0.8	0.8	0.0069
200	10	9.9	10.4	10.4	10.4	0.0158
	5	5.0	5.1	5.1	5.1	0.0162
	1	0.8	0.8	0.8	0.8	0.0052

Note: (1) denotes $\Pr[IM_N \geq u]$, (2) denotes $\Pr[\widehat{IM_N} \geq u]$, (3) denotes $\Pr[IM_N^+ \geq u]$, (4) denotes $\Pr[IM_N \geq \tilde{u}]$ and (5) denotes $|u - \tilde{u}|/u$; here, $\nu_s = 10.5$.

6. Some further issues

This section discusses a few further issues related to the finite-sample corrections considered in this paper. The first issue is: Does $\delta := |u - \tilde{u}|/u$ provide an indication as to when the correction to the critical value of the test can be successfully applied? The values for this quantity corresponding to the nominal levels of 10%, 5% and 1% are given in the final column of Tables 1 through 4. Figure 1 plots the size distortions (in absolute value) of the test based on the corrected critical value against δ corresponding to $\alpha = 5\%$ for all four tests: skewness, excess kurtosis, normality and heteroskedasticity.

Figure 1 shows that for the skewness test, the size-correction can only be effective when $\delta < 0.05$, for the skewness and normality tests the size correction becomes useful when $\delta < 0.1$, and finally the size-correction is effective in the test for heteroskedasticity even for values of δ that exceed 0.3.

The second issue this section addresses is the power of the tests. The heteroskedasticity test is used to provide a comparison of the power of the original test and its finite-sample corrected versions. The data are generated from a normal distribution with $\sigma_i^2 = \exp\{0.5x_i + 0.5x_i^2\}$. Since none of the tests displayed estimated sizes significantly greater than the nominal levels, all tests were based on tabulated critical values. That is, the simulations compare the powers of level α (and not size α) tests. The results based on 10,000 replications are presented in Table 5. Entries are percentages.

The figures in Table 5 show that the corrected tests based on $\widehat{IM}$ and on $\tilde{u}$ displayed in general higher power than the other two tests, that is, the tests based on IM and IM^+.

The second power simulation involves the skewness test and its corrected versions. The simulation design is as described in Section 4. Here, the number of observations is fixed at 100, and the data are generated from a χ_k^2 distribution, for $k = 5, 10, 30, 50$. The lower k, the more asymmetric the distribution of the error terms. The results based on 10,000 replications are given in Table 6. Entries are percentages.

The figures in Table 6 show that the original test was slightly more powerful at the 1% level whereas the corrected tests tended to be more powerful at the 10% and 5% levels. The differences of the estimated powers were, however, very small.

The final issue this section addresses is the sensitivity of the hybrid Bartlett-corrected test to different values of ν_s. In order to examine such sensitivity I ran 10,000 replications of the heteroskedasticity test with $\delta_1 = \delta_2 = 0$ (size) and $\delta_1 = \delta_2 = 0.5$ (power) for $n = 50, 100$. The random seed was set in such a way that all simulations were based on the same set of random numbers, so that the comparisons were meaningful. ν_s was taken and ηth upper quantile of a χ_2^2 distribution, for $\eta = 0.990, 0.991, 0.992, 0.993, 0.994, 0.995, 0.996, 0.997, 0.998, 0.999$. The estimated sizes and powers of the test based on $\widehat{IM}$ were the same for all values of v_s and $\alpha = 10\%, 5\%, 1\%$. That is, neither the size nor the power of the test seemed to be sensitive to the choice of ν_s for ν_s ranging from 9.21 (0.990 upper quantile of a χ_2^2 distribution) to 13.82 (0.999 upper quantile of a χ_2^2 distribution).

Figure 1: Size Distortions for the 5% Level

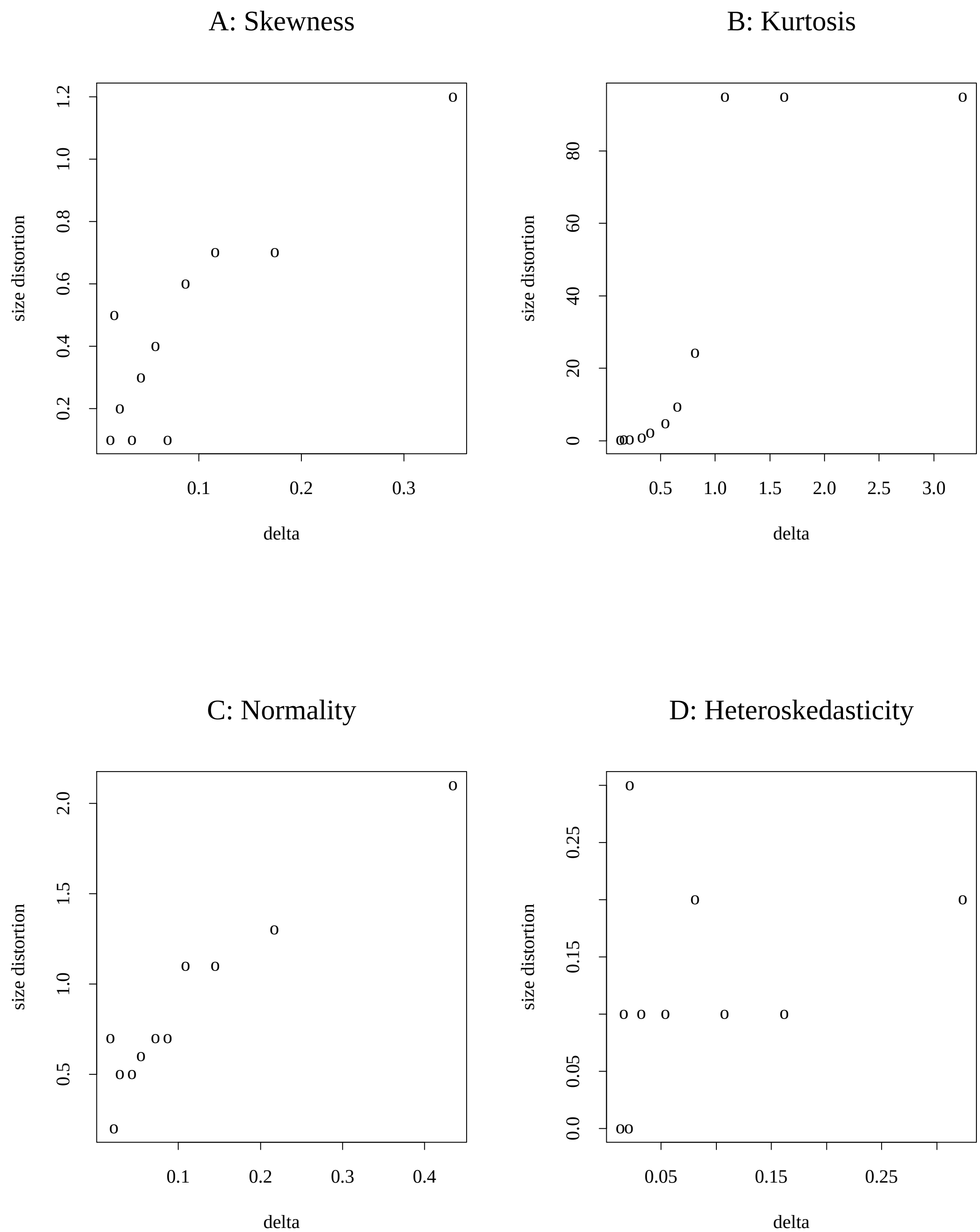

Table 5: Heteroskedasticity Test, Power

n	α	(1)	(2)	(3)	(4)
10	10	10.6	23.5	24.5	28.2
	5	9.9	16.8	17.7	20.2
	1	2.2	5.8	6.1	3.7
20	10	36.9	42.5	42.6	43.9
	5	26.6	32.7	32.4	33.6
	1	11.6	14.3	12.0	13.3
30	10	54.8	58.9	58.6	59.5
	5	42.7	47.9	47.4	48.5
	1	22.6	25.0	21.9	24.3
40	10	68.6	71.2	70.7	71.5
	5	57.5	60.9	59.9	61.2
	1	35.5	37.5	34.0	37.0
50	10	79.0	81.1	80.5	81.3
	5	69.3	71.9	70.7	80.0
	1	47.9	49.4	45.5	49.2
60	10	86.0	87.0	86.1	87.1
	5	78.2	80.2	78.7	80.2
	1	59.1	60.2	55.9	60.1
70	10	91.3	92.1	91.1	92.1
	5	84.8	86.0	84.3	86.1
	1	68.3	69.0	64.3	69.0
80	10	94.2	94.7	93.5	94.7
	5	89.7	90.7	88.5	90.7
	1	75.9	76.8	71.4	76.7
90	10	96.1	96.4	95.0	96.4
	5	93.1	93.6	91.3	93.6
	1	82.2	82.8	77.1	82.8
100	10	97.7	98.0	96.6	98.0
	5	95.6	95.9	93.6	95.9
	1	87.4	87.7	81.6	87.6

Note: (1) denotes $\Pr[IM_N \geq u]$, (2) denotes $\Pr[\widehat{IM_N} \geq u]$, (3) denotes $\Pr[IM_N^{+} \geq u]$, (4) denotes $\Pr[IM_N \geq \tilde{u}]$; here, $\nu_s = 10.5$.

Table 6: Skewness Test, Power

k	α	(1)	(2)	(3)	(4)
50	10	41.6	44.0	44.0	44.0
	5	30.0	31.3	31.4	31.2
	1	14.2	12.5	12.6	13.0
30	10	57.1	59.4	59.5	59.4
	5	45.1	46.5	46.5	46.4
	1	24.6	22.4	22.5	23.0
10	10	93.0	93.8	93.4	93.8
	5	87.2	88.0	87.9	87.9
	1	71.1	68.4	68.4	69.0
5	10	99.4	99.5	99.3	99.5
	5	98.6	98.7	98.4	98.7
	1	94.2	93.0	91.9	93.3

Note: (1) denotes $\Pr[IM_N \geq u]$, (2) denotes $\Pr[\widehat{IM_N} \geq u]$, (3) denotes $\Pr[IM_N^{+} \geq u]$, (4) denotes $\Pr[IM_N \geq \tilde{u}]$; here, $\nu_s = 10.5$.

7. Concluding remarks

This paper argued that it is possible to use the results in Chesher and Spady (1991) to propose a Bartlett-type correction to information matrix test statistics. The test based on this corrected statistic and the test based on transformed critical values proposed by these authors are equivalent to order n^{-1}, *i.e.*, when terms of order smaller than n^{-1} are neglected. The corrections work quite well for some tests. In some cases, they only work properly for moderately large sample sizes. However, in these cases the unmodified test has a dreadful behavior even with large samples. Overall, the results in this paper suggest that if a finite-sample correction to a information matrix test is to be used, the hybrid Bartlett-correction is to be preferred. This corrected test displayed better size and power properties than the competing corrections. Its disadvantage is its clear ad hoc nature. A comforting finding however was that it was found not to be very sensitive to the truncation parameter. Also, the results in this paper suggest that corrections to heteroskedasticity tests tend to be more reliable than corrections to normality tests, despite the fact that the latter are covariate independent. The performance of the finite-sample corrections is particularly poor when it comes to testing for excess kurtosis. Further research should concentrate on explaining why this happens, and designing more reliable corrections.

ACKNOWLEDGEMENTS

I have benefited from conversations with Andrew Chesher, Roger Koenker and Richard Spady. Comments and suggestions from two anonymous referees and Silvia Ferrari which led to a much improved paper are also gratefully acknowledged. This work was partially funded by CNPq/Brazil.

REFERENCES

Breusch, T.S. & Pagan, A.R. (1979). A simple test for heteroskedasticity and random coefficient variation. *Econometrica*, **47**, 1287-1294.

Chesher, A. (1984). Testing for neglected heterogeneity. *Econometrica*, **52**, 865-872.

Chesher, A. & Spady, R. (1991). Asymptotic expansions of the information matrix test statistic. *Econometrica*, **59**, 787-815.

Cordeiro, G.M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265-274.

Cordeiro, G.M. & Ferrari, S.L.P. (1991). A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika*, **78**, 573-582.

Cribari–Neto, F. & Ferrari, S.L.P. (1995). An improved Lagrange multiplier test for heteroskedasticity. *Communications in Statistics—Simulation and Computation*, **24**, 31-44.

Davidson, R. & MacKinnon, J.G. (1992). A new form of the information matrix test. *Econometrica*, **60**, 145-157.

Ferrari, S.L.P. & Cordeiro, G.M. (1994). Matrix formulae for computing improved score tests. *Journal of Statistical Computation and Simulation*, **49**, 195-206.

Fraser, D.A.S. (1965). On information in statistics. *Annals of Mathematical Statistics*, **36**, 890-896.

Godfrey, L.G. (1978). Testing for multiplicative heteroskedasticity. *Journal of Econometrics*, **8**, 227-236.

Harris, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653-659.

Harris, P. (1987). Correction to 'An asymptotic expansion for the null distribution of the efficient score statistic'. *Biometrika*, **74**, 667.

Hill, G.W. & Davis, A.W. (1968). Generalized asymptotic expansions of the Cornish-Fisher type. *Annals of Mathematical Statistics*, **39**, 1264-1273.

Honda, Y. (1988). A size correction to the Lagrange multiplier test for heteroskedasticity. *Journal of Econometrics*, **38**, 375-386.

Jarque, C.M. & Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, **6**, 255-259.

Jarque, C.M. & Bera, A.K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, **55**, 163-172.

Koenker, R. (1981). A note on Studentizing a test for heteroscedasticity. *Journal of Econometrics*, **17**, 107-112.

Orme, C. (1990). The small-sample performance of the information matrix test. *Journal of Econometrics*, **46**, 309-331.

Patel, J.K., Kapadia, C.H. & Owen, D.B. (1976). *Handbook of Statistical Distributions.* New York: Marcel Dekker.

Taylor, L.W. (1987). The size bias of White's information matrix test. *Economics Letters*, **24**, 63-67.

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

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, **48**, 817-838.