Faculdade de Economia
da Universidade de Coimbra

Grupo de Estudos Monetários e Financeiros
(GEMF)
Av. Dias da Silva, 165 – 3004-512 COIMBRA,
PORTUGAL

gemf@fe.uc.pt
http://gemf.fe.uc.pt

JOSÉ MURTEIRA, ESMERALDA RAMALHO
& JOAQUIM RAMALHO

Heteroskedasticity Testing Through Comparison of Wald-Type Statistics

ESTUDOS DO GEMF

N.º 05 2011

PUBLICAÇÃO CO-FINANCIADA PELA
FUNDAÇÃO PARA A CIÊNCIA E TECNOLOGIA

Impresso na Secção de Textos da FEUC
COIMBRA 2011

Heteroskedasticity Testing Through Comparison of Wald-type Statistics*

José M. R. Murteira

Faculdade de Economia, Universidade de Coimbra and CEMAPRE

Esmeralda A. Ramalho and *Joaquim J.S. Ramalho*

Departamento de Economia and CEFAGE-UE, Universidade de Évora

This version: 04/01/2011

* Address for correspondence: José Murteira, Faculdade de Economia, Universidade de Coimbra, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal. E-mail: jmurt@fe.uc.pt.

Abstract

A test for heteroskedasticity within the context of classical linear regression can be based on the difference between Wald statistics in heteroskedasticity-robust and nonrobust forms. The resulting statistic is asymptotically distributed under the null hypothesis of homoskedasticity as chi-squared with one degree of freedom. The power of this test is sensitive to the choice of parametric restriction on which the Wald statistics are based, so the supremum of a range of individual test statistics is proposed. Two versions of a supremum-based test are considered: the first version, easier to implement, does not have a known asymptotic null distribution, so the bootstrap is employed in order to assess its behaviour and enable meaningful conclusions from its use in applied work. The second version has a known asymptotic distribution and, in some cases, is asymptotically pivotal under the null. A small simulation study illustrates the implementation and finite-sample performance of both versions of the test.

JEL classification code: C12, C21.

Key Words: Heteroskedasticity testing; White test; Wald test; Supremum.

Introduction

When testing for homoskedasticity in the context of classical regression, researchers often lack information about the structure of the conditional variance of the dependent variable. A number of tests in the literature can be gathered within a unifying approach, under which homoskedasticity is nested in a continuous skedastic function of a linear combination of regressors functions. Such is the case, *e.g.*, of the well known Glejser (1969) and Godfrey (1978)/Breusch-Pagan (1979) tests, either in their original versions or with subsequent robustness and small sample improvements, as proposed by Koenker (1981), Godfrey (1996), Godfrey and Orme (1999), Machado and Santos Silva (2000) or Im (2000).

Testing for homoskedasticity against a specific alternative is advantageous if the latter coincides with the data generating process (DGP) in case of heteroskedasticity. However, given the frequent lack of information about the variables causing variance heterogeneity, a pure significance test of conditional homoskedasticity may be preferable to more oriented procedures. In this respect, the White (1980) test clearly constitutes the benchmark of an approach that assumes no formal structure about the skedastic process.

As shown by Godfrey and Orme (1999), the fact that the White's test can use many degrees of freedom (df), even for parsimonious models, can have undesirable consequences for the test size and power in small samples. Consequently, it seems useful to try and devise testing procedures more conserving on df's. One possibility is to impose constraints on the coefficients of the artificial regression given in White (1980, eq. 2), *e.g.*, excluding squares and cross-products from this regression. Or, for instance, a test with one df can be obtained by replacing White's regressors with the squared predicted value of the dependent variable (Anscombe, 1961).

As shown below, a heteroskedasticity test with one df also results by considering the difference between Wald-type statistics for restrictions on regression parameters, in heteroskedasticity-robust and nonrobust forms.(1) In line with the results of Godfrey (1996, Appendix 1), the performance of this test is found to be sensitive to the choice of parametric restriction on which the Wald statistics are based. Like all procedures that entail a reduction of the number of df's used by the White's test, the approach incurs the risk of loss of generality relative to the latter and, *e.g.*, the loss of consistency against some heteroskedastic alternatives.

This loss of generality can be attenuated if one takes, as test statistic, the supremum of several tests from a range of different parametric restrictions. In what follows, two versions of this supremum-based approach are presented: the first version, easy to implement through artificial OLS regressions, does not have a known asymptotic null distribution, so the bootstrap is employed in order to assess its behaviour and enable meaningful conclusions from its use in applied work. The second version has a known asymptotic distribution and, in some cases, is asymptotically pivotal under the null. However, as illustrated in a brief Monte Carlo exercise, its asymptotic distribution constitutes a poor approximation to the test distribution in finite samples, so the bootstrap should also be used in this case. This small simulation study indicates that, in some situations, the first version of the supremum-based procedure can outperform conventional tests, including the White's test.

1 Model and Notation

The regression model is $y_i = x_i'\beta + \varepsilon_i$, $i = 1, \ldots, n$, where $\{(x_i', \varepsilon_i), i = 1, \ldots, n\}$ denotes a sequence of independent not necessarily identically distributed (i.n.i.d.) random vectors, such that x_i ($k \times 1$, $k < n$) and the scalar ε_i verify $E(x_i'\varepsilon_i) = 0$. The variables y_i and x_i are observable, while the error term, ε_i, is not. β denotes a $k \times 1$ vector of unknown parameters to be estimated. In this setting, conditional heteroskedasticity is allowed for, generally expressed as

$$E\left(\varepsilon_i^2 | x_i\right) = \sigma^2 \omega\left(x_i'\right) \equiv \sigma^2 \omega_i, \;\; \omega_i > 0, \;\; i = 1, \ldots, n, \tag{1}$$

with $\sigma^2 > 0$ and $\omega(x_i')$ denoting an unspecified, possibly parametric, skedastic function of x_i. It is assumed that the sequence $\{(x_i', \varepsilon_i), i = 1, \ldots, n\}$ satisfies regularity conditions that permit the application of standard asymptotic theory. In particular, assumptions of the type given in White (1980) are adopted throughout the present paper. In matrix notation, (1) can be written as $E(\varepsilon\varepsilon'|X) = \sigma^2 diag(\omega_i, i = 1, \ldots, n) \equiv \sigma^2\Omega$, where $\varepsilon \equiv (\varepsilon_1, \ldots, \varepsilon_n)'$ and X is the conventional $n \times k$ full rank matrix of observations on the vector of covariates, x. As a convenient normalization, let $p\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n}\omega_i = 1$.

Let b denote the OLS estimator of β, providing residuals $e_i \equiv y_i - x_i'b$, $i = 1, \ldots, n$. The usual (homoskedasticity-valid) and heteroskedasticity-robust covariance matrix estimators for b are denoted, respectively, by

$$\begin{aligned}\widehat{V}_1 &\equiv s^2\left(\sum_{i=1}^{n} x_i x_i'\right)^{-1} = s^2(X'X)^{-1},\\ \widehat{V}_2 &\equiv \left(\sum_{i=1}^{n} x_i x_i'\right)^{-1}\left(\sum_{i=1}^{n} e_i^2 x_i x_i'\right)\left(\sum_{i=1}^{n} x_i x_i'\right)^{-1} = (X'X)^{-1}(X'D_eX)(X'X)^{-1},\end{aligned}$$

with D_e denoting an $n \times n$ diagonal matrix with typical diagonal element e_i^2, $i = 1, ..., n$, and $s^2 \equiv n^{-1}\sum_{i=1}^{n} e_i^2$.

The White's test is a test of the null hypothesis (H_0) that consists of the nonredundant restrictions of

$$p\lim_{n\longrightarrow\infty} n\left(\widehat{V}_1 - \widehat{V}_2\right) = 0,$$

equivalent, under standard assumptions, to $p\lim_{n\longrightarrow\infty} n^{-1}\sum_{i=1}^{n}(e_i^2 - s^2)x_i x_i' = 0$. White's "direct test for heteroskedasticity" is obtained as nR^2 from the regression of e_i^2 on a constant term and the nonredundant terms in $x_i x_i'$, where R^2 denotes the usual coefficient of determination. As is well known, this statistic is asymptotically distributed under H_0 as chi-squared with, at most, $k(k+1)/2$ df's.

Next, consider a vector function, $r(\beta)$, where $r(\cdot) : \mathbb{R}^k \to \mathbb{R}^j$ denotes a vector of $j\ (< k)$ functionally independent, continuously differentiable, functions of β. The set of $j \times 1$ vector of restrictions, $r(\beta) = 0$, will henceforth be termed *auxiliary restriction*. Let $R(\beta) \equiv \partial r(\beta)/\partial\beta'$, the $j \times k$ Jacobian of $r(\beta)$ with respect to β. Functional independence in $r(\beta)$ ensures full row rank of $R(\beta)$ for all β.

Define the $n \times j$ matrix $T \equiv X(X'X)^{-1}R(b)'$. Then, the Wald statistics associated with the test of the auxiliary restriction, in nonrobust (W_{NR}) and robust (W_R)

forms can be written, respectively, as

$$
\begin{aligned}
W_{NR} &\equiv r(b)' \left[R(b) \widehat{V}_1 R(b)' \right]^{-1} r(b) = r(b)' \left(s^2 T'T \right)^{-1} r(b), \\
W_R &\equiv r(b)' \left[R(b) \widehat{V}_2 R(b)' \right]^{-1} r(b) = r(b)' \left(T'D_e T \right)^{-1} r(b).
\end{aligned}
$$

The following definitions will also be used in the ensuing text:

$$
\begin{aligned}
\underset{(k\times k)}{\Sigma_X} &\equiv p \lim_{n\to\infty} \left(n^{-1} X'X \right), \\
\underset{(k\times k)}{\Xi_X} &\equiv p \lim_{n\to\infty} \left(n^{-1} X' D_e X \right) = \sigma^2 p \lim_{n\to\infty} \left(n^{-1} X' \Omega X \right),
\end{aligned}
$$

with existence of probability limits ensured by White's (1980) Assumptions 2 and 3, and the last equality shown by White (1980, Theorem 1)]. The $j \times j$ matrices Σ_T and Ξ_T are analogously defined, with T replacing X. In addition consider

$$
M_n \equiv n^{-1} X'X = n^{-1} \sum_{i=1}^{n} x_i x_i',
$$

$$
\underset{(k\times 1)}{\gamma_1} \equiv \Sigma_X^{-1} R(\beta)' \Sigma_T^{-1} r(\beta), \quad \underset{(k\times 1)}{\gamma_2} \equiv \Sigma_X^{-1} R(\beta)' \Xi_T^{-1} r(\beta), \tag{2}
$$

$$
\underset{(k\times 1)}{c_1(b)} \equiv \left(X'X \right)^{-1} R(b)' \left(T'T \right)^{-1} r(b), \tag{3}
$$

$$
\underset{(k\times 1)}{c_2(b)} \equiv \left(X'X \right)^{-1} R(b)' \left(T'D_e T \right)^{-1} r(b),
$$

$$
\alpha_i \equiv \gamma_1' x_i x_i' \gamma_2, \quad \bar{\alpha} \equiv n^{-1} \sum_{i=1}^{n} \alpha_i = \gamma_1' M_n \gamma_2,
$$

$$
a_i \equiv c_1(b)' x_i x_i' c_2(b), \quad \bar{a} = n^{-1} \sum_{i=1}^{n} a_i = c_1(b)' M_n c_2(b).
$$

2 Difference Between Wald Statistics

The following Lemma can be established:

Lemma 1 *If the auxiliary restriction is false, that is, $r(\beta) \neq 0$ at the true value of the regression parameters, then, under Assumptions 1, 2(b), 3(b), 5-7 in White (1980), if ε_i is independent of x_i and homoskedastic, $\forall i$,*

$$
\frac{1}{v} \left[n^{-1/2} s^2 \left(W_R - W_{NR} \right) \right]^2 \xrightarrow{D} \chi_1^2,
$$

where $\upsilon \equiv n^{-1} \sum_{i=1}^{n} E\left[\left(\sigma^2 - \varepsilon_i^2\right)^2 (\alpha_i - \bar{\alpha})^2\right]$ *and* χ_1^2 *denotes the chi-squared distribution with one df.*

A feasible statistic can be obtained by replacing υ with the consistent estmator

$$\upsilon \equiv n^{-1} \sum_{i=1}^{n} \left[\left(s^2 - e_i^2\right)(a_i - \bar{a})\right]^2.$$

Then, a test statistic, asymptotically distributed under H_0 as a chi-squared random variable (rv) with one df, results as

$$\frac{\left[s^2 \left(W_R - W_{NR}\right)\right]^2}{\sum_{i=1}^{n} \left[\left(s^2 - e_i^2\right)(a_i - \bar{a})\right]^2}. \tag{4}$$

If, in Lemma 1, Assumption 7 of White (1980) is replaced with the assumption that the ε_i are homokurtic, $\forall i$ $\left[E\left(\varepsilon_i^4\right) = \mu_4, \forall i\right]$, the test can be performed through a simplified procedure, as stated in the next Remark.

Remark 1 *If* $r(\beta) \neq 0$, *a "direct test" of* H_0 *can be obtained, as in White (1980, eq. 2), from the OLS regression* $e_i^2 = \zeta_0 + \zeta_1 a_i + residuals$. *Under Assumptions 1, 2(b), 3(b), 5 and 6 in White (1980), if* ε_i *is independent of* x_i, *homoskedastic and homokurtic,* $\forall i$, *a procedure that is asymptotically equivalent to the test that results from (4) is the test of* $\gamma_1 = 0$ *using the standard* R^2 *statistic from this regression. Formally,*

$$nR^2 \xrightarrow{D} \chi_1^2. \tag{5}$$

Special cases of interest of the above results are stated as Corollaries.

Corollary 1 *If* $r(\cdot)$ *is a scalar function* ($j = 1$), *then (4) and (5) are valid test statistics, whether* $r(\beta) = 0$ *is true or false.*

When scalar affine auxiliary restrictions are employed, further results can be obtained, enabling computation of the test through simplified procedures using common econometrics packages.

Corollary 2 *If $r(\cdot)$ is a scalar affine function, write $\theta \equiv r(\beta) = R\beta - r$, with R a row k-vector of constants and r a scalar; then*

(i) The test statistics (4) and (5) are asymptotically pivotal.

(ii) Let $\theta = R_1\beta_1 + R_2\beta_2 - r$, where R and β are partitioned into conformable ($k-1$)-vectors (R_1 and β_1) and scalars (R_2 and β_2); let x_i be conformably partitioned as $\left(x_{i1}' \vdots\, x_{i2}\right)'$ and let $x_{i1}^* \equiv x_{i1} - (x_{i2}/R_2) R_1'$; then, the statistic referred to in (5) can also be computed as nR^2 from the OLS regression

$$e_i^2 = \zeta_0 + \zeta_1 u_i^2 + residuals, \tag{6}$$

where u_i denotes the i-th OLS residual from the regression of x_{i2} on x_{i1}^*.

The proposed test is consistent whenever heteroskedasticity causes the two versions of the Wald-type statistic to diverge. Specifically, this approach tests the significance of $n^{-1/2}\sum_{i=1}^{n}(s^2 - e_i^2)\,a_i$, which, under standard assumptions, is $O_p\left(n^{-1/2}\right)$ if $n^{-1}\sum_{i=1}^{n} x_i x_i' \varepsilon_i^2$ and $\sigma^2 M_n$ are not asymptotically equivalent. The following Lemma presents the asymptotic distribution of the test under a sequence of local alternative hypotheses.

Lemma 2 *Under the sequence of local alternatives $H_1 : E(\varepsilon_i^2|x_i) = \sigma^2\omega\left(n^{-1/2}z_i'\eta\right)$, with z_i and η denoting l-vectors of, respectively, functions of x_i and unknown parameters, $\omega(0) \equiv \left[\omega\left(n^{-1/2}z_i'\eta\right)\right]_{\eta=0} = 1$ and $\omega'(0) \equiv \left[d\omega_i/d\left(n^{-1/2}z_i'\eta\right)\right]_{\eta=0} \neq 0$,*

$$\frac{1}{\upsilon}\left[n^{-1/2}s^2\left(W_R - W_{NR}\right)\right]^2 \xrightarrow{D} \chi_1^2(\lambda), \quad \lambda \equiv \mu^2/\upsilon,$$

where

$$\mu \equiv -\sigma^2\omega'(0)\lim_{n\to\infty} E\left[(\alpha_i - \bar{\alpha})\left(n^{-1/2}z_i'\eta\right)\right]$$

and $\chi_1^2(\lambda)$ denotes the noncentral chi-squared distribution with one df and noncentrality parameter λ.

The numerator in the noncentrality parameter can be seen to increase (decrease) as the covariance between z_i and α_i increases (decreases) in absolute value. This indicates that the choice of $r(\beta)$ can affect the performance of the test in finite samples. Ideally, $r(\beta)$ should be selected so as to achieve a high value of λ in case of heteroskedasticity. However, this can obviously be difficult, in view of the frequent lack of information about the structure of heteroskedasticity.(2)

3 Supremum of Differences Between Wald Statistics

The sensitivity of the test performance to the particular auxiliary restriction may be attenuated if one uses as test statistic the supremum of different statistics [from either (4), (5) or (6)], obtained from a range of parametric restrictions. Presumably, the supremum of such a range is positively influenced by the more powerful tests against the unknown skedastic alternative, which tend to produce higher statistics. Let this test be named "sup-r test".

Clearly, the statistics from particular auxiliary restrictions are not independent under H_0, which makes it difficult to obtain the null distribution of the supremum. Therefore, the bootstrap should be used, so as to approximate this distribution and to perform the sup-r test. Alternatively, one can consider the supremum of orthogonalised statistics, whose limit null distribution can be established, due to asymptotic independence. To this effect, consider, first, m auxiliary restrictions $r_g(\beta) = 0$, $g = 1, ..., m$, and corresponding robust and nonrobust Wald statistics, $W_R^{(g)}$, $W_{NR}^{(g)}$, and define $\alpha_i^{(g)}$ and $\overline{\alpha^{(g)}}$ analogously as, respectively, α_i and $\overline{\alpha}$, for each auxiliary restriction $r_g(\beta) = 0$. The next Lemma constitutes the basis for a modified version of the sup-r test.

Lemma 3 *Let* $wd \equiv \begin{pmatrix} W_R^{(1)} - W_{NR}^{(1)} & W_R^{(2)} - W_{NR}^{(2)} & \cdots & W_R^{(m)} - W_{NR}^{(m)} \end{pmatrix}'$*, the m-vector of Wald statistics differences; assume that the functions $r_g(\cdot)$ are functionally independent and that $r_g(\beta) \neq 0$, $g = 1, ..., m$ at the true value of the parameters β. Define the $m \times m$ matrix Υ with typical element*

$$n^{-1} \sum_{i=1}^{n} E\left\{ \left(\sigma^2 - \varepsilon_i^2\right)^2 \left[\alpha_i^{(g)} - \overline{\alpha^{(g)}}\right] \left[\alpha_i^{(h)} - \overline{\alpha^{(h)}}\right] \right\}, \quad g, h = 1, \ldots m. \tag{7}$$

Let the symmetric positive definite (pd) matrix $\Psi \equiv \Upsilon^{-1/2}$ *denote the square root of the matrix* Υ^{-1}. *Then, under Assumptions 1, 2(b), 3(b), 5-7 in White (1980), if* ε_i *is independent of* x_i *and homoskedastic,* $\forall i$,

$$\Psi \times n^{-1/2} s^2 wd \xrightarrow{D} N(0_m, I_m), \tag{8}$$

where $N(0_m, I_m)$ *denotes the* m*-variate standard normal distribution (with* 0_m *a null* m*-vector and* I_m *the identity matrix of order* m*).*

Lemma 3 implies that the standardized Wald statistics differences are asymptotically independent under homoskedasticity. From this result one can obtain the asymptotic distribution of the supremum of those differences, as formally stated in the next Corollary.

Corollary 3 *Partition* Ψ *into its* m *column vectors,* $\Psi = \begin{bmatrix} \Psi_1 & \cdots & \Psi_m \end{bmatrix}$. *Under* H_0 *and White's (1980) Assumptions,*

$$n^{-1} s^4 \sup\left\{ (\Psi_1' wd)^2, \ \ldots, \ (\Psi_m' wd)^2 \right\} \xrightarrow{D} C_m,$$

where C_m *denotes the chi-squared distribution with one df, raised to power* m.

The average covariance matrix Υ can be estimated by the matrix V with elements

$$V_{gh} \equiv n^{-1} \sum_{i=1}^{n} \left(s^2 - e_i^2\right)^2 \left[a_i^{(g)} - \overline{a^{(g)}}\right] \left[a_i^{(h)} - \overline{a^{(h)}}\right], \quad g, h = 1, \ldots m,$$

with $a_i^{(g)}$ defined analogously as a_i, for each auxiliary restriction $r_g(\beta) = 0$, $g = 1, ..., m$. Given the continuity of the square root function, defined on the set of positive definite matrices (see, *e.g.*, Horn and Johnson, 1999, Ch. 7.2), the elements of Ψ can be estimated by the corresponding elements of the (matrix) square root of V^{-1} (name it P). Partition P as $\begin{bmatrix} P_1 & \cdots & P_m \end{bmatrix}$; the statistics obtained by replacing Ψ with P in (8) are asymptotically independent normal, so

$$n^{-1} s^4 \sup\left\{ (P_1' wd)^2, \ldots, (P_m' wd)^2 \right\} \tag{9}$$

constitutes a feasible test statistic corresponding to the rv in Corollary 3.

Let swd denote the observed value of this version of the sup-r statistic and let $C_m^{-1}(\xi)$, $\xi \in (0,1)$, denote the $\xi \times 100\%$ quantile of the chi-squared distribution with one df. Then, H_0 is rejected at the $\alpha \times 100\%$ nominal significance level if $swd > C_m^{-1}\left[(1-\alpha)^{1/m}\right]$.

The following Corollaries are analogous to Corollaries 1 and 2(i) above:

Corollary 4 *If the functions $r_g(\cdot)$, $g = 1, ..., m$, are scalars, then $Pn^{-1/2}s^2wd$ only depends on β through the values of $R_g(\beta)$ [not directly through the values of $r_g(\beta)$, $g = 1, ..., m$].*

Corollary 5 *If the functions $r_g(\cdot)$, $g = 1, ..., m$, are scalar affine, then, under H_0,*

$$n^{-1}s^4 \sup\left\{(P_1'wd)^2, \ \ldots, \ (P_m'wd)^2\right\}$$

is an asymptotically pivotal statistic.

Given the result of Beran (1988) on the use of the bootstrap with asymptotically pivotal statistics, the bootstrap can be employed here in conjunction with m scalar affine auxiliary restrictions, so as to achieve more reliable control over the performance of this version of the sup-r test in finite samples. Meanwhile, the statistics referred to in (4), (5) or (6) are not (even asymptotically) independent for different auxiliary restrictions, under H_0. As is well known, for dependent rv's t_1, ..., t_m,

$$\Pr\left(\sup\{t_1, ..., t_m\} \leq t\right) = \Pr\left(t_1 \leq t, ..., t_m \leq t\right) \neq \prod_{g=1}^{m} \Pr\left(t_g \leq t\right),$$

which raises the issue of the dependence structure of the t_g, upon which their joint distribution also depends. Thus, the null distribution of the supremum of statistics from (4), (5) or (6) is not invariant to the type of dependence among individual tests, which means that the corresponding test statistic is not asymptotically pivotal. Thus, even though the bootstrap can be employed in conjunction with these statistics, it does not yield an asymptotic refinement, when compared with first-order asymptotic approximation results.

4 Monte Carlo Illustration

A brief simulation exercise now illustrates the implementation and behaviour of the proposed tests. The data are generated by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \;\; i = 1, \ldots, n, \tag{10}$$

with parameters set to one and regressors obtained as independent random vectors from a bivariate normal distribution with zero mean vector, unit marginal variances, and correlation 0.65. The disturbances ε_i are iid draws from one of the following distributions: standard normal, $N(0,1)$, Student's t with five df's, t_5, and chi-squared with two df's, χ_2^2. In each case ε_i is transformed to have zero mean and one of the following conditional variances:

Homoskedasticity: $H_0 : V(\varepsilon_i | x_{i1}, x_{i2}) = 1.$

Heteroskedasticity: $H_1 : V(\varepsilon_i | x_{i1}, x_{i2}) = (1 + 4x_{i2}^2)/5.$

$H_2 : V(\varepsilon_i | x_{i1}, x_{i2}) = \left[1 + \sum_{l=1}^{2} (4 - l)\, x_{il}^2\right]/6.$

$H_3 : V(\varepsilon_i | x_{i1}, x_{i2}) = \exp(x_{i1} + x_{i2} - 1.65).$

Under H_1 and H_2 the conditional variance is specified as in Machado and Santos Silva (2000); both specifications result from random variation of the slope coefficients, a frequent cause for heteroskedasticity in empirical applications. Under H_1, $V(\beta_2) = 4$ and, under H_2, $V(\beta_l) = 4 - l$, $l =$ 1,2, with different weights attributed to x_1 and x_2. Under H_3 the skedastic function depends on regressors levels, rather than their squares. In all cases $E[V(\varepsilon_i | x_{i1}, x_{i2})] = 1$.

The following tests are considered: the "studentized" form of the Breusch-Pagan test, due to Koenker(1981) (denoted as B-P/K); the White's test (W); a test computed as nR^2 from the regression of e_i^2 on an intercept and the square of the dependent variable fitted value (Anscombe, 1961) (A); two tests based on the difference between Wald statistics for each of the following scalar affine auxiliary restrictions: $r_1(\beta) \equiv \beta_0 + \beta_1 + \beta_2 = 0$ (r_1) and $r_2(\beta) \equiv \beta_1 + \beta_2 = 0$ (r_2); a test based on the difference between Wald statistics for the joint auxiliary restriction $r(\beta) = 0$, where $r(\beta) \equiv [r_1(\beta), r_2(\beta)]'$ (r_c); and, finally, two forms of the sup-r test, based on the auxiliary restrictions $r_1(\beta) = 0$ and $r_2(\beta) = 0$ [sup-r_A – supremum of nR^2 statistics from the r_1 and r_2 tests; sup-r_B – test based on the statistic sup $\left\{ (P_1'wd)^2, (P_2'wd)^2 \right\}$, with $wd = \left(\begin{array}{cc} W_R^{(1)} - W_{NR}^{(1)} & W_R^{(2)} - W_{NR}^{(2)} \end{array} \right)'$ and P_1 and P_2 as defined in (9)]. The statistics B-P/K, W, r_1, r_2 and sup-r_B are asymptotically pivotal under the null hypothesis.

The test denoted as r_c is computed as nR^2 from the regression referred to in Remark 1. It is noted that, as required by Lemma 1, the artificial restriction $r(\beta) = 0$ is false. The r_1 and r_2 tests are computed as nR^2 from (6); the corresponding auxiliary restrictions $r_g(\beta) \equiv \theta = 0$, $g = 1, 2$, yield the following reparameterizations of model (10):

$$r_1 : y_i = \beta_1 (1 - x_{i2}) + \beta_2 (x_{i1} - x_{i2}) + \theta x_{i2} + \varepsilon_i; \qquad r_2 : y_i = \beta_1 + \beta_2 (x_{i1} - x_{i2}) + \theta x_{i2} + \varepsilon_i.$$

Then, the term u_i in (6) denotes the OLS residual from the regression of x_{i2} on, respectively,

$$x_{i1}^* \equiv (1 - x_{i2}, x_{i1} - x_{i2})', \; [r_1(\beta) = 0]; \qquad x_{i1}^* \equiv (1, x_{i1} - x_{i2})', \; [r_2(\beta) = 0].$$

Tables 1 and 2 contain percentages of rejections for the eight tests at the 5% nominal significance level, based on 10000 replications of samples with size $n = 100$ and with regressors newly drawn at each replication.(3) Results in Table 1 estimate the size of the tests, both from asymptotic and bootstrap critical values. Following Hodoshima and Ando (2007), the nonparametric residual bootstrap is used, with 499 bootstrap resamples and residuals in each bootstrap resample multiplied by $\sqrt{n/(n-3)}$.(4) An asterisk flags cases for which 5% lies outside a 95% confidence interval for the true rejection probability of the null. Computations were performed with TSP v.4.5 (Hall and Cummins, 1999).

The bootstrap seems to provide better control over the significance level than asymptotic theory in several cases of asymptotically pivotal Koenker-type tests [namely, W with t_5 and χ_2^2 errors, r_1 and r_2 with $N(0,1)$ and $t(5)$ errors]. This appears to be in line with Beran (1988) as well as the results and reccomendations of Godfrey and Orme (1999) and Godfrey, Orme and Santos Silva (2006) on the use of the nonparametric bootstrap for such tests. Results for the A test also indicate a better performance of the bootstrap [with $N(0,1)$ and t_5 errors]. Under all null error distributions this is also the case for the r_c test and, especially, the sup-r_B test, found to severely overreject the null on the basis of critical values from the asymptotic distribution (C_2). The null asymptotic distribution of the sup-r_A test is not known so the bootstrap is used in this case (a simulation-based approach is not useful, because the error distribution is supposed unknown by the researcher).

Table 1 – Percentage of Rejections at the 5% Nominal Level, under Homoskedasticity

Error Distribution	$N(0,1)$		t_5		χ_2^2	
Test	asy	boot	asy	boot	asy	boot
B-P/K	4.91	5.27	4.89	5.22	6.09*	5.59*
W	5.40	5.60*	6.88*	5.68*	8.50*	6.34*
A	4.41*	5.09	4.50*	5.31	5.01	5.40
r_1	4.41*	5.12	4.27*	5.09	4.93	5.21
r_2	4.42*	5.41	4.50*	5.43*	4.63	5.47*
r_c	4.20*	5.17	3.82*	5.10	4.00*	5.04
sup-r_A	–	5.31	–	5.28	–	5.52*
sup-r_B	9.36*	5.33	7.38*	4.34*	7.39*	4.22*

*: 5% rejection probability outside 95% confidence interval.
Values refer to either asymptotic critical values (columns "asy") or bootstrap critical values (columns "boot").

Table 2 presents estimates of the probability of rejection of the null hypothesis under H_1 through H_3. All percentages are computed with reference to bootstrap-based critical values: although size estimates in Table 1 do not afford a clear-cut choice, this option seems preferable to using asymptotic critical values in the majority of cases considered in the exercise.

Even within a succint study such as the present one, the low power of the sup-r_B test is noteworthy. Use of the sup-r_A version seems clearly preferable, competing in equal terms with conventional tests under H_1 (W) and H_3 (B-P/K and A), and outperforming them in the remaining cases. The rejection percentages for this test are positively influenced by the most powerful of r_1 and r_2 tests, the performance of which (in line with theoretical predictions) looks quite sensitive to the particular form of heteroskedasticity. It is interesting to note the contrast between the power of the sup-r_A test and that of the r_c test, which appears to be attracted by the least powerful of r_1 and r_2 tests (or performs even worse than either of these, under H_3). The sup-r_A procedure thus seems the best choice among the different tests involving differences between Wald-type statistics and, quite often, among all the tests considered in the exercise.

5 Concluding Remark

The approach proposed in the present paper yields a test that, according to a limited simulation study, seems to compete rather well with existing tests for heteroskedasticity. The study is merely illustrative and, naturally, begs the question of the test behaviour under more general circumstances. Meanwhile, the present methodology suggests some topics for future research, including, among others, the use of the proposed procedure within the general framework of the information matrix test.

Table 2 – Percentage of Rejections at the 5% Nominal Level under Heteroskedasticity

Error Distribution	$N(0,1)$	t_5	χ^2_2
H_1: $V\left(\varepsilon_i \vert x_{i1}, x_{i2}\right) = \left(1+4x_{i2}^2\right)/5$			
B-P/K	27.66	24.36	24.38
W	92.91	73.98	61.96
A	74.69	59.29	50.00
r_1	56.41	44.69	38.78
r_2	90.81	76.01	67.29
r_c	58.68	46.67	40.75
sup-r_A	88.06	72.76	62.20
sup-r_B	18.50	12.15	13.87
H_2: $V\left(\varepsilon_i \vert x_{i1}, x_{i2}\right) = [1 + \sum_{l=1}^{2} (5-l)\, x_{il}^2]/6$			
B-P/K	28.53	25.50	24.73
W	91.67	72.24	58.12
A	81.80	64.46	53.54
r_1	62.30	48.94	40.76
r_2	95.36	82.19	72.44
r_c	64.42	50.56	42.80
sup-r_A	93.90	78.37	67.08
sup-r_B	24.70	15.03	17.14
H_3: $V\left(\varepsilon_i \vert x_{i1}, x_{i2}\right) = \exp(x_{i1} + x_{i2}$ - 1.65)			
B-P/K	99.98	98.11	95.89
W	99.29	92.83	86.77
A	99.93	98.71	97.77
r_1	99.98	99.19	98.36
r_2	89.43	79.21	74.72
r_c	82.36	75.96	74.42
sup-r_A	99.94	98.35	96.62
sup-r_B	70.58	57.82	50.36

6 Proofs

Proof of Lemma 1. The scaled difference between Wald statistics can be successively written as

$$n^{-1/2}s^2\left(W_R - W_{NR}\right) = n^{-1/2}r\left(b\right)'\left[s^2\left(T'D_eT\right)^{-1} - \left(T'T\right)^{-1}\right]r\left(b\right) =$$
$$n^{-1/2}r\left(b\right)'\left(T'T\right)^{-1}T'\left(s^2I_n - D_e\right)T\left(T'D_eT\right)^{-1}r\left(b\right) =$$
$$c_1\left(b\right)'\left[n^{-1/2}X'\left(s^2I_n - D_e\right)X\right]c_2\left(b\right) = c_1\left(b\right)'\left[n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)x_ix_i'\right]c_2\left(b\right) =$$
$$c_1\left(b\right)'\left[n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)\left(x_ix_i' - M_n\right)\right]c_2\left(b\right).$$

Generally speaking, under White's Assumptions, $\Sigma_X = n^{-1}X'X + O_p\left(n^{-1/2}\right)$ and $\Xi_X = n^{-1}X'D_eX + O_p\left(n^{-1/2}\right)$. Then,

$$\Sigma_T \equiv p\lim_{n\to\infty}\left(n^{-1}T'T\right) = R\left(\beta\right)\Sigma_X^{-1}R\left(\beta\right)' = n^{-1}T'T + O_p\left(n^{-1/2}\right),$$
$$\Xi_T \equiv p\lim_{n\to\infty}\left(n^{-1}T'D_eT\right) = R\left(\beta\right)\Sigma_X^{-1}\Xi_X\Sigma_X^{-1}R\left(\beta\right)' = n^{-1}T'D_eT + O_p\left(n^{-1/2}\right).$$

Also, from the definitions of γ_j and $c_j\left(b\right)$, $j = 1, 2$ [in (2) and (3), respectively], $c_j\left(b\right) = \gamma_j + O_p\left(n^{-1/2}\right)$. Under H_0 and White's Assumptions,

$$n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)x_ix_i' = O_p\left(1\right)$$

(White, 1980, Theorem 2). From this result, one can write

$$n^{-1/2}s^2\left(W_R - W_{NR}\right) = \gamma_1'\left[n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)\left(x_ix_i' - M_n\right)\right]\gamma_2 + o_p\left(1\right) = \delta_n + o_p\left(1\right),$$

where

$$\delta_n \equiv \gamma_1'\left[n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)\left(x_ix_i' - M_n\right)\right]\gamma_2 = n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)\left(\alpha_i - \bar{\alpha}\right).$$

White (1980, Theorem 2) shows that, under homoskedasticity, the elements of $n^{-1/2}\sum_{i=1}^{n}\left(s^2 - e_i^2\right)\left(x_ix_i' - M_n\right)$ have limit normal distributions. Thus, provided that $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$ [implying $r\left(\beta\right) \neq 0$], under H_0, $\delta_n \xrightarrow{D} N\left(0, \lim_{n\to\infty} \upsilon\right)$, with

$$\upsilon \equiv n^{-1}\sum_{i=1}^{n} E\left\{\left[\left(\sigma^2-\varepsilon_i^2\right)\gamma_1'\left(x_ix_i'-M_n\right)\gamma_2\right]^2\right\} = n^{-1}\sum_{i=1}^{n} E\left\{\left[\left(\sigma^2-\varepsilon_i^2\right)\left(\alpha_i-\bar{\alpha}\right)\right]^2\right\}. \tag{11}$$

Replacing γ_j with $c_j(b)$, $j=1,2$, and recalling the definition of a_i, it immediately follows that

$$n^{-1/2}s^2\left(W_R-W_{NR}\right) = n^{-1/2}\sum_{i=1}^{n}\left(s^2-e_i^2\right)a_i$$

$$= n^{-1/2}\sum_{i=1}^{n}\left(s^2-e_i^2\right)\left(a_i-\bar{a}\right) = \delta_n + o_p(1),$$

so $n^{-1/2}s^2\left(W_R-W_{NR}\right) \xrightarrow{D} N\left(0,\lim_{n\to\infty}\upsilon\right)$ as well. Then the required result immediately follows. ■

Proof of Remark 1. Remark 1 immediately follows from Corollary 1 of White (1980). ■

Proof of Corollary 1. If $r(\cdot)$ is a scalar function, then T is a column $n-$vector and $a_i = r(b)^2 T_{(i)}^2 / \left(T'TT'D_eT\right)$, where $T_{(i)} \equiv x_i'\left(X'X\right)^{-1}R(b)'$ denotes the i-th element of T. Let $\overline{T^2} \equiv n^{-1}\sum_{i=1}^{n} T_{(i)}^2$; cancelling out constant terms, the statistic in (4) becomes

$$\frac{\left[n^{-1/2}s^2\left(W_R-W_{NR}\right)\right]^2}{\upsilon} = \frac{\left[\sum_{i=1}^{n}\left(s^2-e_i^2\right)a_i\right]^2}{\sum_{i=1}^{n}\left(s^2-\varepsilon_i^2\right)^2\left(a_i-\bar{a}\right)^2} = \frac{\left[\sum_{i=1}^{n}\left(s^2-e_i^2\right)T_{(i)}^2\right]^2}{\sum_{i=1}^{n}\left(s^2-e_i^2\right)^2\left(T_{(i)}^2-\overline{T^2}\right)^2},$$

which does not involve $r(b)$.

The direct test can be obtained through the artificial regression of e_i^2 on a constant term and the regressor $T_{(i)}^2$, which is just a_i rescaled. Now, if e denotes the n-vector of OLS residuals, a common result has $e = M\varepsilon$, with $M \equiv I_n - X\left(X'X\right)^{-1}X'$, not involving β. As e_i – and, hence, s^2 – do not involve β, and $T_{(i)}$ only depends on β through $R(b)$, both statistics, in (4) and (5), converge in distribution to the chi-squared distribution with one df, regardless of whether $r(\beta)$ is zero or not [obviously, $r(\cdot)$ should be differentiable in the parameter space of interest]. ■

Proof of Corollary 2. (i) If $r(\beta) = R\beta - r$, a scalar, then $\partial r(\beta)/\partial\beta' = R$, a vector of constants not involving β. Thus, from Corollary 1, the test statistic no longer depends on β and, consequently, it is asymptotically pivotal. (It is not pivotal, because its finite sample distribution depends upon the error distribution.)

(ii) The result is a direct consequence of the fact that u_i^2 in (6) is proportional to $\left[x_i'(X'X)^{-1}R'\right]^2$ when $r(\cdot)$ is a scalar affine function. To see this, start by writing the reparameterized model in matrix form as $y^* = X^*\beta^* + \varepsilon$, where

$$X \equiv \left[\underset{n\times(k-1)}{X_1} \,\vdots\, \underset{n\times 1}{x_2}\right], \quad x_2^* \equiv (1/R_2)\,x_2, \quad y^* \equiv y - x_2^*, \quad \beta^* \equiv (\beta_1', \theta)',$$

$$X^* \equiv \left[X_1^* \,\vdots\, x_2^*\right] = XA, \quad \underset{k\times k}{A} = \begin{bmatrix} I_{k-1} & 0 \\ -(1/R_2)\,R_1 & 1/R_2 \end{bmatrix},$$

with θ and R defined in the main text and I_{k-1} denoting the identity matrix of order $k-1$. Under the reparameterized model the auxiliary restriction becomes $\theta = R^*\beta^* = 0$, $R^* \equiv RA$.

The direct test can be computed as nR^2 from the OLS regression of e_i^2 on an intercept and the regressor $\left[x_i'(X'X)^{-1}R'\right]^2$. In matrix form, the n-vector with generic element $x_i'(X'X)^{-1}R'$ can be written $X(X'X)^{-1}R'$. As A is invertible,

$$X(X'X)^{-1}R' = XA(A'X'XA)^{-1}A'R' = X^*(X^{*\prime}X^*)^{-1}R^{*\prime}.$$

The residuals from the original and reparameterized model (e^*) are equal, because

$$e^* = M^*y^* = My^* = My - (1/R_2)\,Mx_2 = My = e,$$

where $M^* = I - X^*(X^{*\prime}X^*)^{-1}X^{*\prime} = I - X(X'X)^{-1}X' = M$ and $Mx_2 = 0$, since M projects onto the space orthogonal to the space spanned by the columns of X. Thus, the direct test can also be computed as nR^2 from the OLS regression of e_i^{*2} on an intercept and the regressor $\left[x_i^{*\prime}(X^{*\prime}X^*)^{-1}R^{*\prime}\right]^2$. From the definition of R^* and the usual formulae for the inverse of partitioned matrices, the n-vector with generic element $x_i^{*\prime}(X^{*\prime}X^*)^{-1}R^{*\prime}$ is given by

$$X^*(X^{*\prime}X^*)^{-1}R^{*\prime} = (x_2^{*\prime}M_1^*x_2^*)^{-1}M_1^*x_2^*,$$

where $M_1^* \equiv I_{k-1} - X_1^* \left(X_1^{*\prime} X_1^* \right)^{-1} X_1^{*\prime}$ and $X_1^* \equiv X_1 - x_2^* R_1$. This is proportional to the vector of OLS residuals from the regression of x_2^* on X_1^*, proportional, in turn, to $M_1^* x_2$, the n-vector of OLS residuals from the regression of x_2 on X_1^*. Thus, $\left[x_i^{*\prime} \left(X^{*\prime} X^* \right)^{-1} R^{*\prime} \right]^2$ and u_i^2 are proportional, so the regression of e_i^2 on an intercept and $\left[x_i' \left(X'X \right)^{-1} R' \right]^2$ and regression (6) yield the same nR^2 statistic. ■

Proof of Lemma 2. With $\omega(0) = 1$, a first-order Taylor expansion of ω_i around $\eta = 0$ leads to

$$\omega\left(n^{-1/2} z_i' \eta\right) \simeq \omega(0) + \omega'(0)\, n^{-1/2} z_i' \eta = 1 + \omega'(0)\, n^{-1/2} z_i' \eta.$$

Thus, under H_1 and White's Assumptions, the elements of $n^{-1/2} \sum_{i=1}^n \left(s^2 - e_i^2\right) \left(x_i x_i' - M_n\right)$ are asymptotically normal with means given by the elements of

$$\lim_{n\to\infty} E\left[\left(\sigma^2 - \varepsilon_i^2\right)\left(x_i x_i' - M_n\right)\right] = -\sigma^2 \omega'(0) \lim_{n\to\infty} E\left[\left(x_i x_i' - M_n\right) n^{-1/2} z_i' \eta\right].$$

It immediately follows that $\delta_n \xrightarrow{D} N(\mu, \upsilon)$, with

$$\mu \equiv -\sigma^2 \omega'(0) \lim_{n\to\infty} E\left[\left(\alpha_i - \bar{\alpha}\right)\left(n^{-1/2} z_i' \eta\right)\right]$$

and υ defined in (11). Obviously, then, $n^{-1/2} s^2 \left(W_R - W_{NR}\right) \xrightarrow{D} N(\mu, \upsilon)$ as well. Standard results from Statistics ensure that, under the sequence H_1,

$$\frac{1}{\upsilon}\left[n^{-1/2} s^2 \left(W_R - W_{NR}\right)\right]^2 \xrightarrow{D} \chi_1^2(\lambda),$$

with noncentrality parameter $\lambda \equiv \mu^2/\upsilon$. ■

Proof of Lemma 3. For each auxiliary restriction, $r_g(\beta) = 0$, $g = 1, ..., m$, write the corresponding element of the vector $n^{-1/2} s^2 w d$ as

$$n^{-1/2} s^2 \left[W_R^{(g)} - W_{NR}^{(g)}\right] = \delta_n^{(g)} + o_p(1),$$

where (in obvious notation)

$$\delta_n^{(g)} \equiv n^{-1/2} \sum_{i=1}^n \left(s^2 - e_i^2\right) \gamma_1^{(g)\prime} \left(x_i x_i' - M_n\right) \gamma_2^{(g)} = n^{-1/2} \sum_{i=1}^n \left(s^2 - e_i^2\right) \left[\alpha_i^{(g)} - \overline{\alpha^{(g)}}\right].$$

Under White's (1980) Assumptions and with homoskedastic errors independent of x_i, the multivariate Liapounov central limit theorem can be applied to the random m-vector $n^{-1/2}\left(\delta_n^{(1)},\ldots,\delta_n^{(m)}\right)'$, that is,

$$\Psi \times n^{-1/2}\left(\delta_n^{(1)},\ldots,\delta_n^{(m)}\right)' \xrightarrow{D} N\left(0_m, I_m\right),$$

where the symmetric pd matrix Ψ is such, that $\Psi^2 = \Upsilon^{-1}$ and Υ is the average covariance matrix defined in (7). The existence of Ψ is ensured by the independence of x_i and ε_i and White's Assumptions 5 and 6, guaranteeing that Υ is a pd matrix with uniformly bounded elements for sufficiently large n. The asymptotic equivalence between each $n^{-1/2}s^2\left[W_R^{(g)} - W_{NR}^{(g)}\right]$ and $\delta_n^{(g)}$ then yields the statement in the present Lemma. ■

Proof of Corollary 3. Consider the components of the $m\times 1$ vector $\Psi n^{-1/2}s^2 wd$,

$$\left(\Psi_1' n^{-1/2}s^2 wd, \ldots, \Psi_m' n^{-1/2}s^2 wd\right)'.$$

According to Lemma 3, these components are asymptotically uncorrelated standard normal rv's, so they are asymptotically independent. Thus, the corresponding squared variables, $\left(\Psi_g' n^{-1/2}s^2 wd\right)^2$, are asymptotically independent chi-squared with one df. The desired result immediately follows from the well-known fact that, for independent rv's t_g, $g = 1, ..., m$,

$$\Pr\left(\sup\left\{t_1, ..., t_m\right\} \leq t\right) = \Pr\left(t_1 \leq t, ..., t_m \leq t\right) = \prod\nolimits_{g=1}^{m} \Pr\left(t_g \leq t\right).$$

■

Proof of Corollary 4. If the functions $r_g(\cdot)$ are all scalar, then $a_i^{(g)} = r_g(b)^2 T_{g(i)}^2 / \left(T_g' T_g T_g' D_e T_g\right)$ [where all quantities are defined with reference to the auxiliary restriction $r_g(\beta) = 0$, analogously as before with reference to $r(\beta) = 0$]. Thus,

$$s^2\left[W_R^{(g)} - W_{NR}^{(g)}\right] = r_g(b)^2 / \left(T_g' T_g T_g' D_e T_g\right) \sum\nolimits_{i=1}^{n} \left(s^2 - e_i^2\right)^2 T_{g(i)}^2.$$

Define the $m \times m$ diagonal matrix $DR(b)$, with $r_g(b)^2 / \left(T_g' T_g T_g' D_e T_g\right)$ as g-th diagonal entry. Then, the matrix V can be written as

$$V = DR(b) \times M \times DR(b), \tag{12}$$

where M is $m \times m$ symmetric and pd for large enough n, with generic element

$$M_{gh} \equiv \sum\nolimits_{i=1}^{n} \left(s^2 - e_i^2\right)^2 \left[T_{g(i)}^2 - \overline{T_g^2}\right] \left[T_{h(i)}^2 - \overline{T_h^2}\right], \quad g, h = 1, ..., m.$$

Given the definitions of T_g and $T_{g(i)}$, $g = 1, ..., m$, it should be stressed that M depends on β only through the derivatives of the functions $r_g(b)$, $R_g(b) \equiv \partial r_g(b) / \partial b'$ – not through the functions $r_g(\cdot)$, $g = 1, ..., m$.

From (12),

$$V^{-1} = DR(b)^{-1} \times M^{-1} \times DR(b)^{-1},$$

where, obviously, $DR(b)^{-1}$ is diagonal with g-th entry $\left(T_g' T_g T_g' D_e T_g\right) / r_g(b)^2$, $g = 1, ..., m$. Thus, considering the symmetric square root matrix of M^{-1} (denoted as PM),

$$V^{-1} = DR(b)^{-1} \times PM^2 \times DR(b)^{-1} = \left[DR(b)^{-1} PM\right]^2 = P^2,$$

from which $P = DR(b)^{-1} PM$, where PM depends on β only through the derivatives $R_g(b)$, $g = 1, ..., m$.

Then, finally, $P \times wd = PM \times \left[DR(b)^{-1} wd\right]$, which is an m-vector depending on the auxiliary restrictions only through $R_g(b)$, $g = 1, ..., m$: this is because the functions $r_g(b)$ are canceled out in $DR(b)^{-1} wd$. Thus, when all the functions $r_g(\cdot)$, $g = 1, ..., m$, are scalars, the vector $Pn^{-1/2} s^2 wd$ does not depend directly on the value of the $r_g(\beta)$. ■

Proof of Corollary 5. If $r_g(\beta) = R_g \beta - q$, then $R_g(b) = R_g$, $g = 1, ..., m$ vectors of constants not involving b. Thus, M (and PM) are independent of b, the only link of $P \times wd$ to β. Therefore $P \times wd$ is a vector of asymptotically pivotal statistics. From Lemma 3, these statistics are asymptotically independent.

For independent rv's, t_g, $g = 1, ..., m$, $\Pr(\sup\{t_1, ..., t_m\} \leq v) = \prod_{g=1}^{m} \Pr(t_g \leq t)$. Thus, if every t_g is asymptotically pivotal for all DGP's in H_0, then each $\Pr(t_g \leq t)$, $g = 1, ..., m$ – and so $\Pr(\sup\{t_1, ..., t_m\} \leq t)$ – is invariant under all DGP's in H_0. Therefore, $\sup\{P_1' wd, \ldots, P_m' wd\}$ is asymptotically pivotal because the $P_g' wd$, $g = 1, ..., m$, are asymptotically pivotal independent statistics. Obviously, this statement applies to $\sup\left\{(P_1' wd)^2, \ldots, (P_m' wd)^2\right\}$ as well. ■

Notes

(1) The idea is remotely inspired by the Hausman (1978) test, as applied to a test statistic contrast rather than an estimator difference.

(2) Results allow an interpretation of the test as a check of the impact of heteroskedasticity on inferences about specific parameter restrictions. Failure to reject the null leads to the conclusion that heteroskedasticity, if present, does not affect W_{NR} significantly. If a particular restriction is of interest, using the test with that restriction can be useful. If the null is not rejected, then inference on that restriction may proceed with the nonrobust covariance estimate.

(3) At 1% and 10% levels results follow similar patterns, so they are omitted.

(4) This is "boot1" method in Hodoshima and Ando (2007). With the White's "direct" test under homoskedasticity, the approach is found by the authors to work best, overall, among other bootstrap methods (including variants of the wild bootstrap of Mammen, 1993, or Davidson and Flachaire, 2008).

References

Anscombe, F. (1961), "Examination of Residuals", *Proceedings from the Berkeley Symposium*, 1, 1-36.

Beran, R. (1988), "Prepivoting Test Statistics: A Bootstrap View of Asymptotic Refinements", *Journal of the American Statistical Association*, 83, 687-697.

Breusch, T. S., A. R. Pagan (1979), "A Simple Test for Heteroskedasticity and Random Coefficients Variation", *Econometrica*, 47, 1287-1294.

Davidson, R., E. Flachaire (2008), "The Wild Bootstrap, Tamed at Last", *Journal of Econometrics*, 146(1), 162-169.

Glejser, H. (1969), "A New Test for Heteroskedasticity", *Journal of the American Statistical Association*, 64, 316-323.

Godfrey, L. G. (1978), "Testing for Multiplicative Heteroskedasticity", *Journal of Econometrics*, 8, 227-236.

Godfrey, L. G. (1996), "Some Results on the Glejser and Koenker Tests for Heteroskedasticity", *Journal of Econometrics*, 72, 275-299.

Godfrey, L. G., C. D. Orme (1999), "The Robustness, Reliability and Power of Heteroskedasticity Tests", *Econometric Reviews*, 18(2), 169-194.

Godfrey, L. G., C. D. Orme, J. M. C. Santos Silva (2006), "Simulation-based Tests for Heteroskedasticity in Linear Regression Models: Some Further Results", *Econometrics Journal*, 9, 76-97.

Hall, B. H., C. Cummins (1999), *TSP User's Guide, Version 4.5*, TSP International, Palo Alto, Ca.

Hausman, J. A. (1978), "Specification Tests in Econometrics", *Econometrica*, 46, 1251-1272.

Hodoshima, J., M. Ando (2007), "The Finite-Sample Performance of White's Test for Heteroskedasticity Under Stochastic Regressors", *Communications in Statistics – Simulation and Computation*, 36, 1201–1215.

Horn, R. A., C. R. Johnson (1990), *Matrix Analysis*, Cambridge University Press.

Im, K. S. (2000), "Robustifying the Glejser Test of Heteroskedasticity", *Journal of Econometrics*, 97, 179-188.

Koenker, R. (1981), "A Note on Studentizing a Test for Heteroskedasticity", *Journal of Econometrics*, 17, 107-112.

Machado, J. A. F., J. M. C. Santos Silva (2000), "Glejser's Test Revisited", *Journal of Econometrics*, 97, 189-202.

Mammen, E. (1993), "Bootstrap and Wild Bootstrap for High Dimensional Linear Models", *Annals of Statistics*, 21, 255–285.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, *48*, 817-838.

ESTUDOS DO G.E.M.F.

(Available on-line at http://gemf.fe.uc.pt)

2011-05 *Heteroskedasticity Testing Through Comparison of Wald-Type Statistics*
- José Murteira, Esmeralda Ramalho & Joaquim Ramalho

2011-04 *Accession to the European Union, Interest Rates and Indebtedness: Greece and Portugal*
- Pedro Bação & António Portugal Duarte

2011-03 *Economic Voting in Portuguese Municipal Elections*
- Rodrigo Martins & Francisco José Veiga

2011-02 *Application of a structural model to a wholesale electricity market: The Spanish market from January 1999 to June 2007*
- Vítor Marques, Adelino Fortunato & Isabel Soares

2011-01 *A Smoothed-Distribution Form of Nadaraya-Watson Estimation*
- Ralph W. Bailey & John T. Addison

2010-22 *Business Survival in Portuguese Regions*
- Alcina Nunes & Elsa de Morais Sarmento

2010-21 *A Closer Look at the World Business Cycle Synchronization*
- Pedro André Cerqueira

2010-20 *Does Schumpeterian Creative Destruction Lead to Higher Productivity? The effects of firms' entry*
- Carlos Carreira & Paulino Teixeira

2010-19 *How Do Central Banks React to Wealth Composition and Asset Prices?*
- Vítor Castro & Ricardo M. Sousa

2010-18 *The duration of business cycle expansions and contractions: Are there change-points in duration dependence?*
- Vítor Castro

2010-17 *Water Pricing and Social Equity in Portuguese Municipalities*
- Rita Martins, Carlota Quintal, Eduardo Barata & Luís Cruz

2010-16 *Financial constraints: Are there differences between manufacturing and services?*
- Filipe Silva & Carlos Carreira

2010-15 *Measuring firms' financial constraints: Evidence for Portugal through different approaches*
- Filipe Silva & Carlos Carreira

2010-14 *Exchange Rate Target Zones: A Survey of the Literature*
- António Portugal Duarte, João Sousa Andrade & Adelaide Duarte

2010-13 *Is foreign trade important for regional growth? Empirical evidence from Portugal*
- Elias Soukiazis & Micaela Antunes

2010-12 *MCMC, likelihood estimation and identifiability problems in DLM models*
- António Alberto Santos

2010-11 *Regional growth in Portugal: assessing the contribution of earnings and education inequality*
- Adelaide Duarte & Marta Simões

2010-10 *Business Demography Dynamics in Portugal: A Semi-Parametric Survival Analysis*
- Alcina Nunes & Elsa Sarmento

2010-09 *Business Demography Dynamics in Portugal: A Non-Parametric Survival Analysis*
- Alcina Nunes & Elsa Sarmento

2010-08 *The impact of EU integration on the Portuguese distribution of employees' earnings*
- João A. S. Andrade, Adelaide P. S. Duarte & Marta C. N. Simões

2010-07 *Fiscal sustainability and the accuracy of macroeconomic forecasts: do supranational forecasts rather than government forecasts make a difference?*
- Carlos Fonseca Marinheiro

2010-06 *Estimation of Risk-Neutral Density Surfaces*
- A. M. Monteiro, R. H. Tütüncü & L. N. Vicente

2010-05 *Productivity, wages, and the returns to firm-provided training: who is grabbing the biggest share?*
- Ana Sofia Lopes & Paulino Teixeira

2010-04 *Health Status Determinants in the OECD Countries. A Panel Data Approach with Endogenous Regressors*
- Ana Poças & Elias Soukiazis

2010-03 *Employment, exchange rates and labour market rigidity*
- Fernando Alexandre, Pedro Bação, João Cerejeira & Miguel Portela

2010-02 *Slip Sliding Away: Further Union Decline in Germany and Britain*
- John T. Addison, Alex Bryson, Paulino Teixeira & André Pahnke

2010-01 *The Demand for Excess Reserves in the Euro Area and the Impact of the Current Credit Crisis*
- Fátima Teresa Sol Murta & Ana Margarida Garcia

2009-16 *The performance of the European Stock Markets: a time-varying Sharpe ratio approach*
- José A. Soares da Fonseca

2009-15 *Exchange Rate Mean Reversion within a Target Zone: Evidence from a Country on the Periphery of the ERM*
- António Portugal Duarte, João Sousa Andrade & Adelaide Duarte

2009-14 *The Extent of Collective Bargaining and Workplace Representation: Transitions between States and their Determinants. A Comparative Analysis of Germany and Great Britain*
- John T. Addison, Alex Bryson, Paulino Teixeira, André Pahnke & Lutz Bellmann

2009-13 *How well the balance-of- payments constraint approach explains the Portuguese growth performance. Empirical evidence for the 1965-2008 period*
- Micaela Antunes & Elias Soukiazis

2009-12 *Atypical Work: Who Gets It, and Where Does It Lead? Some U.S. Evidence Using the NLSY79*
- John T. Addison, Chad Cotti & Christopher J. Surfield

2009-11 *The PIGS, does the Group Exist? An empirical macroeconomic analysis based on the Okun Law*
- João Sousa Andrade

2009-10 *A Política Monetária do BCE. Uma estratégia original para a estabilidade nominal*
- João Sousa Andrade

2009-09 *Wage Dispersion in a Partially Unionized Labor Force*
- John T. Addison, Ralph W. Bailey & W. Stanley Siebert

2009-08 *Employment and exchange rates: the role of openness and technology*
- Fernando Alexandre, Pedro Bação, João Cerejeira & Miguel Portela

2009-07 *Channels of transmission of inequality to growth: A survey of the theory and evidence from a Portuguese perspective*
- Adelaide Duarte & Marta Simões

2009-06 *No Deep Pockets: Some stylized results on firms' financial constraints*
- Filipe Silva & Carlos Carreira

2009-05 *Aggregate and sector-specific exchange rate indexes for the Portuguese economy*
- Fernando Alexandre, Pedro Bação, João Cerejeira & Miguel Portela

2009-04 *Rent Seeking at Plant Level: An Application of the Card-De La Rica Tenure Model to Workers in German Works Councils*
- John T. Addison, Paulino Teixeira & Thomas Zwick

2009-03 *Unobserved Worker Ability, Firm Heterogeneity, and the Returns to Schooling and Training*
- Ana Sofia Lopes & Paulino Teixeira

2009-02 *Worker Directors: A German Product that Didn't Export?*
- John T. Addison & Claus Schnabel

2009-01 *Fiscal and Monetary Policies in a Keynesian Stock-flow Consistent Model*
- Edwin Le Heron

A série Estudos do GEMF foi iniciada em 1996.